MECHANISTIC STUDIES OF RADICAL SAM ENZYMES IN CYSTOBACTAMID, ANAEROBILIN AND THIAMIN BIOSYNTHESIS

A Dissertation

by

YUANYOU WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Tadhg P. Begley
Committee Members,	Frank M. Raushel
	Coran M. H. Watanabe
	Craig D. Kaplan
Head of Department,	Simon W. North

May 2019

Major Subject: Chemistry

ABSTRACT

The radical S-adenosylmethionine (SAM) enzymes represent a large class enzyme that catalyzes a wide range of reactions, including chemically challenging reactions in critical pathways across all domains of life. Work here describes the detailed mechanistic studies of three of these enzymes in unusual methylation, ring cleavage, and rearrangement reactions.

CysS is a cobalamin-dependent radical SAM methyltransferase that catalyzes the iterative methylations on the 3-methoxy-4-aminobenzoic acid moieties of the antibiotic cystobactamids. We were able to reconstitute the activity of CysS *in vitro* and demonstrate that it can catalyze sequential methylations from a methyl group to form not only ethyl, isopropyl groups but also *sec*-butyl, *t*-butyl groups by a radical mechanism for the first time. To further elucidate the mechanism, different strategic substrate analogs were designed to trap the organic radical species, cobalamin intermediates and probe the radical substitution step. Based on all the isotope labeling studies and substrate analogs experiments, we were able to build a detailed mechanistic model to reveal how CysS uses cobalamin, SAM, [4Fe− 4S] cluster to perform the unusual iterative methylations.

ChuW can catalyze a radical-mediated reaction to methylate and break the protoporphyrin ring of heme to release iron in an anaerobic heme degradation pathway. We were able to design a cyclopropyl heme analog to trap radical generated on the

porphyrin ring by forming a porphyrin compound with cyclopropyl ring cleaved. Several tetrapyrrole ring and SAM adducts were also observed, which gives us further insight into the mechanism of the enzymatic porphyrin cleavage.

ThiC is a non-canonical radical enzyme that catalyzes one of the most complicated rearrangements in all of the mechanistic enzymology, which is the conversion of the 5-aminoimidazole ribonucleotide (AIR) to form 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) in thiamin biosynthesis. Here, we were able to trap a four-carbon fragment from the ribose part of AIR by methionine during the enzymatic reaction. Using isotope labeling substrates and cofactors, we were able to show the intermediate leaks out after the first hydrogen atom abstraction but before the loss of C3′ of AIR.

DEDICATION

To my parents, who brought me up to be what I am...

ACKNOWLEDGEMENTS

Ph.D. is an exciting journey towards unknown territory. However, it would be impossible to truly enjoy the incredible scenery without the people who inspired and motived me along the trip.

First and foremost, I would like to thank Dr. Tadhg Begley for being a remarkable advisor. It's Dr. Begley who leads me to experience the beauty of the logic of chemistry thinking and its application into complicated biological questions, especially in the arrow pushing enzymatic mechanism and chemical logic in the reconstitution of biosynthesis pathway. I am grateful for the freedom and patience Tadhg gave me to work on some very challenging projects. I not only got the best enzymology and organic chemistry training from Tadhg but also was inspired by his enthusiasm and critical thinking for science during the years.

I would also like to thank my committee members, Dr. Frank Raushel, Dr. Coran Watanabe and Dr. Craig Kaplan, for serving on my graduate committees. I am grateful for their insightful comments and encouragement during my seminar, presentations and preliminary examinations.

My sincere thanks also go to all my collaborators for the three projects I have been worked on. I would like to thank Prof. Rolf Müller Müller, Bastien, Sascha for providing the plasmids for the CysS project, Prof. Catherine Drennan, Jennifer, Percy for the crystallography collaboration, Prof. William Lanzilotta for the protein plasmids and all the discussion for the ChuW project. I also like to thank Prof. Coran Watanabe for

her help in the activation of the peptide carrier protein and Prof. Squire Booker for providing me with the cobalamin uptake plasmid.

During my Ph.D. study, there are many people who give me tremendous support. I would like to thank Sameh for teaching me the “trick” in organic synthesis and how to obtain a “clean” NMR spectrum, Seth and Deepika from Barondeau lab for helping me to run the size exclusion column in their glove box, Dmytro for providing me with initial building blocks for the photoaffinity project, Jian for helping me to make different constructs for CysS, ChuW and microbiology experiments, Berkir for training me on the ITC instrument, Angad for helping me to start with the radical SAM enzymes, Lisa for mentoring me in the initial molecular biology experiments.

Thanks to DJ for helping me troubleshooting during the years and Prem for suggestion from many aspects, Rung-yi and Niklamal for insightful suggestions in my early graduate studies, Yiquan and Yindrila for MS instrument, experiment training, Nitai and Lina for their enthusiasm and encouragements in science, Sanjoy, Isita for their suggestion and discussion in flavin enzymes, Sumedh, for many stimulating discussions and advice on the radical SAM enzymes. I also enjoy the time with the energetic new students Anushree, Avick, Indranil, Saad, Sohan, Hannah in Begley lab.

Last but not the least, I would like to thank my family: my parents and my cousin for both spiritually and physically support and “tolerating” me throughout the graduate study.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Tadhg P. Begley (advisor), Professor Frank M. Raushel, and Professor Coran M. H. Watanabe of the Department of Chemistry and Professor Craig D. Kaplan of the Department of Biochemistry and Biophysics.

The CysS plasmids and the initial small molecules for the substrate screening of CysS were obtained from Professor Rolf Müller in Helmholtz Center for Infection Research. The CysG1 plasmids were provided by Bastien Schnell, Dr. Sascha Baumann in Professor Rolf Müller lab.

The origin ChuW plasmids, information about the substrates specificity of ChuW were provided by Professor William Nicholas Lanzilotta in the University of Georgia.

The labeling AIR substrate analogs were synthesized by Dr. Sameh Abdelwahed.

The crystallography study of CysS is collaborated with Prof. Catherine Drennan's lab.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Chemistry department at Texas A&M University.

This work was also made possible in part by the Robert A. Welch Foundation under Grant Number A-0034 and National Institutes of Health(NIH) under Grant Number DK44083.

NOMENCLATURE

5′-dA	5′-deoxyadenosine
ACP	Acyl carrier protein
ABB	4-amino-3-butoxybenzoic acid
AIR	Aminoimidazole ribotide
AMB	4-amino-3-methoxybenzoate acid
APB	4-amino-3-isopropoxybenzoic acid
B_{12}	Cobalamin
CIP	Calf intestine phosphatase
CO	Carbon monoxide
CoA	Coenzyme A
EPR	Electron paramagnetic resonance
IRN	Imidazole ribonucleotide
HMP-P	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate
HMP	4-amino-5-hydroxymethyl-2-methylpyrimidine
HPLC	High performance liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
NMR	Nuclear magnetic resonance
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth
MeCbl	Methylcobalamin

NADPH	Nicotinamide adenine dinucleotide phosphate hydrate
NRP	Nonribosomal peptide
PK	Polyketides
SAM	S-adenosyl-L-methionine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Tris•HCl	Tris(hydroxymethyl)aminomethane hydrochloride

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

1. INTRODUCTION

1.1. Radical S-Adenosylmethionine (SAM) Enzymes

Radical SAM enzymes, first classified as a superfamily in 2001,[1] utilize a [4Fe−4S] cluster and SAM to initiate a diverse set of radical reactions that are otherwise difficult or impossible to catalyze by polar mechanisms. Radical SAM enzymes usually involve H-atom abstraction from an inactivated C−H bond via generation of a 5′-deoxyadenosyl radical (5′-dA•) intermediate by reductive [4Fe−4S] cluster (Figure 1.1). Compared with iron(IV)-oxo-derived radicals, radical generated by hydrogen abstraction from 5′-dA is more persistent because iron-bound oxygen species is not available.[2] Since the high intrinsic reactivity of organic radicals, radical SAM enzymes can catalyze a wide range of reactions, including oxidation, methylation, double-bond addition, and fragmentation reactions.[3]

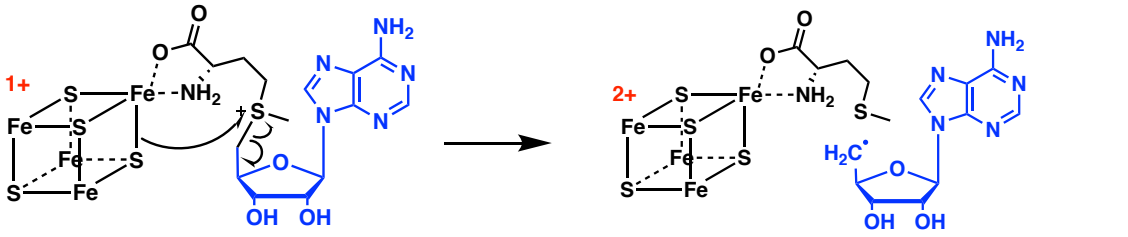


Figure 1.1 Reductive cleavage of SAM to the 5′-dA radical by [4Fe−4S] cluster

For example, many radical SAM enzymes are involved in complicated cofactor biosynthesis (Figure 1.2). The thiamin pyrimidine synthase (ThiC) catalyzes a complex rearrangement of aminoimidazole ribotide (AIR) to form hydroxymethylpyrimidine phosphate (HMP-P) in thiamin biosynthesis.[4] Hydroxybenzimidazole synthase (BzaF) carries out a similar rearrangement of AIR to form 5- hydroxybenzimidazole (HBI) in cobalamin biosynthesis.[5] Aminofutalosine synthase (MqnE) catalyzes the addition of the adenosyl radical to the double bond of 3-[(1-carboxyvinyl)oxy]benzoic acid to form aminofutalosine in the menaquinone biosynthesis.[6-7] F_0-synthase (CofG/CofH or FbiC) cleaves tyrosine and couples it to pyrimidinedione to give the deazaflavin chromophore in coenzyme F_{420} biosynthesis.[8] MoaA and MoaC catalyze the rearrangement of GTP into a cyclic pyranopterin monophosphate (cPMP) in molybdopterin biosynthesis.[9-10]

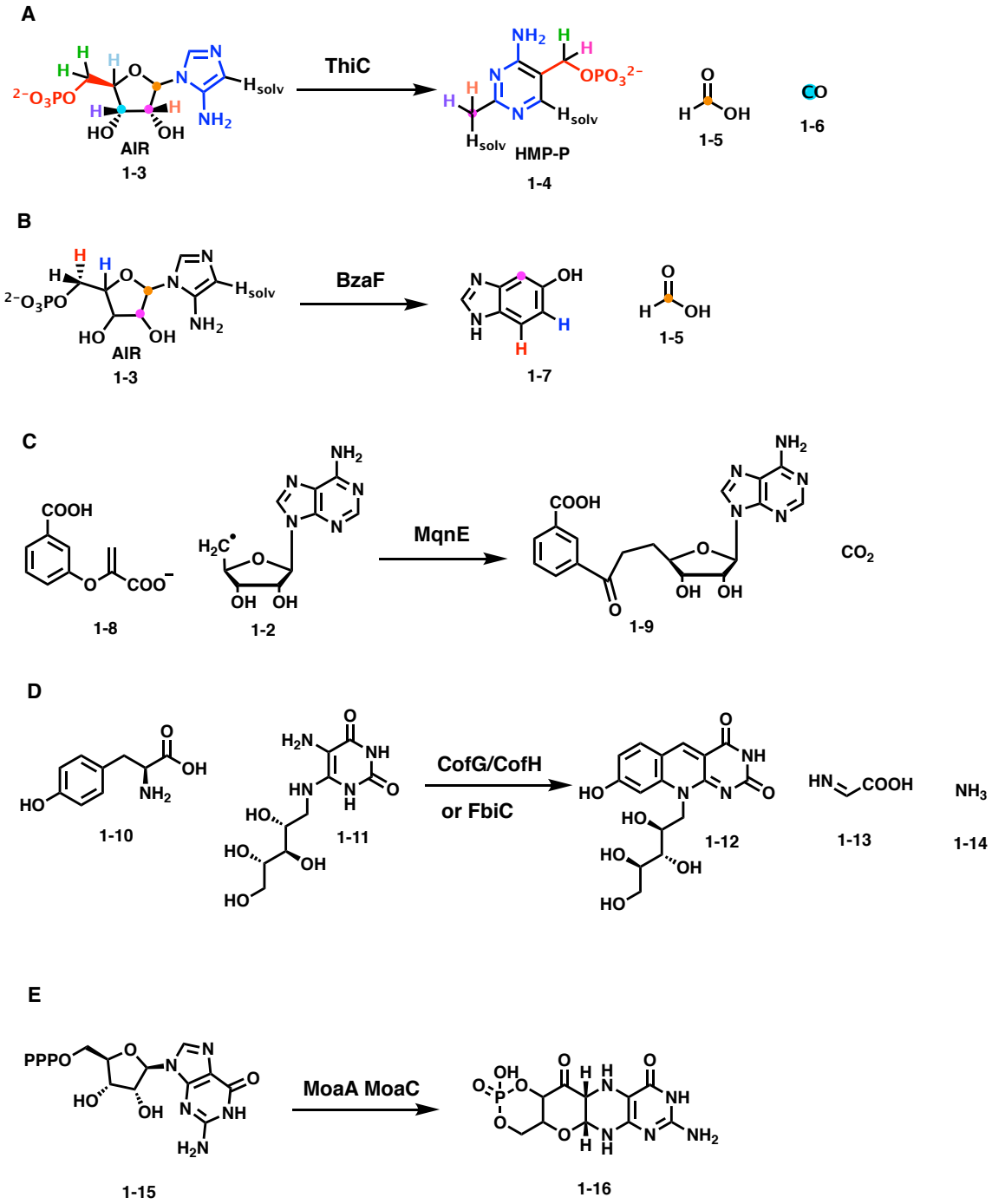


Figure 1.2 Complex radical reactions catalyzed by radical SAM enzymes. (A) ThiC in thiamin biosynthesis (B) BzaF in cobalamin biosynthesis (C) MqnE in menaquinone biosynthesis (D) CofG/CofH in F420 biosynthesis (E) MoaA-MoaC in molybdopterin biosynthesis.

The radical SAM superfamily has grown rapidly since 2001 and now contains almost 113,000 unique sequences that represent more than 65 distinct reactions.[11] Around 50,000 of the sequences relate to proteins have unknown function. Based on the radical SAM enzyme's capability of catalyzing unusual reactions, it's considered as a new treasure for the discovery of novel enzymatic transformations and biochemical pathways.

1.1.1. Non-canonical radical SAM enzymes

For most of the racial SAM enzymes that have been identified, their sequences contain a consensus CX_3CX_2C motif where each of the three cysteines coordinates with one of the three irons in [4Fe–4S] cluster. The fourth iron coordinates with the amino and carboxylate groups of SAM in a bidentate fashion. However, there are several enzymes that do not contain CX_3CX_2C motif but still generate a 5′-deoxyadenosyl radical to catalyze the reactions, such as ThiC (CX_2CX_4C), PhnJ ($CX_2CX_{21}CX_5C$), QueE ($CX_{14}CX_2C$).[4, 12-13]

Besides the difference of the conserved sequence for binding the [4Fe–4S] cluster, ThiC also lacks the classic SAM binding motifs.[14] From the ThiC crystal structure with SAH or 5′-dA and L-methionine, it is suggested that in ThiC active site, SAM anchors to an additional active site metal through its amino and carboxylate groups instead of the iron in [4Fe–4S] cluster.[15] It's still not clear whether the captured conformation is in the catalytic cycle and whether this extra iron helps ThiC to catalyze such complicated rearrangement. What's more, the [4Fe–4S] cluster-binding loop of ThiC is at the dimer interface of the C-terminal instead of N-terminal for the classic

radical enzymes. This swapped structure is most resembled the adenosylcobalamin (AdoCbl)-dependent enzymes (like glutamate mutase), which indicates an evolutionary relationship between the two super-families.

Dph2 is another non-canonical radical SAM enzyme that uses reduced [4Fe-4S] cluster and SAM, but generates a 3-amino-3-carboxypropyl radical instead of a 5′-deoxyadenosyl radical in the diphthamide biosynthesis. It is shown that Dph2 adopts a completely unrelated tertiary structure for [4Fe-4S] cluster binding and $C_{\gamma,Met}$ of SAM is closer to the unique Fe than the S_{Met} in the classical radical SAM enzymes.[16] This structure orientation helps to explain why SAM is reductively cleaved to form 3-amino-3-carboxypropyl radical instead of a 5′-deoxyadenosyl radical.

1.1.2. Radical SAM methyltransferases

Before radical SAM enzyme has been classified as a superfamily, SAM was best known as a natural methyl donor in many important biological processes, including regulation of gene expression, protein modification, RNA processing.[17] In classical methyltransferase (MTase) reactions, substrates need strong nucleophiles such as oxygen, nitrogen or sulfur to attack the methyl group of SAM, which is electrophilic due to the adjacent sulfonium of SAM.[18] However, this nucleophilic S_N2 methylation mechanism cannot apply to the saturated carbon atoms or electrophilic substrates. Alternatively, nature uses radical mechanisms to catalyze these unusual reactions.

To date, all the known enzymes that catalyze radical-mediated methylation belong to the radical SAM superfamily. Sequence analysis reveals that the currently identified radical SAM methyltransferases (RSMTs) can be grouped into several classes

(Figure 1.3).[19-20] Class A RSMTs have a single conserved radical SAM domain and use two conserved cysteine residues to methylate sp^2-hybridized carbon atoms. Class B RSMTs include a cobalamin binding domain (CBD) at the N-terminal as well as a radical SAM domain at the C-terminus. Class C RSMTs has high sequence similarity with HemN at the C-terminus in addition to the classic TIM barrel radical SAM domain. Currently, RlmN and Cfr in the class A RSMTs are the most well-studied methyltransferases in radical SAM family.[20] There are more and more reconstitutions and mechanistic studies for Class B RSMTs while there are few reconstituted examples for Class C RSMTs.

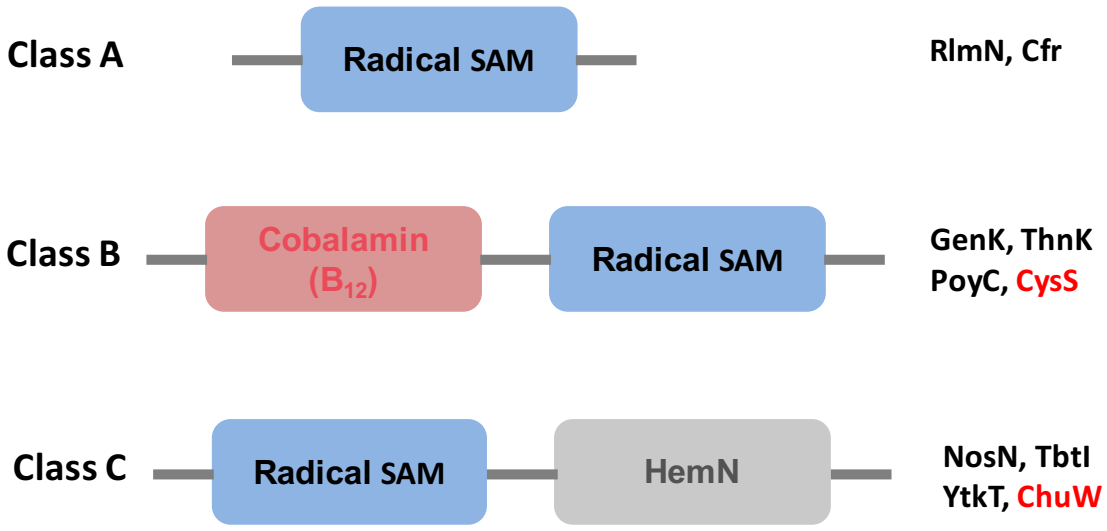


Figure 1.3 Domain structures of different classes of RSMTs.

As the data shown in 2017, the number of total radical SAM sequences are nearly 255 000. Among them, around 7000 radical SAM proteins contain CBDs based on

bioinformatics.[21-22] Currently, class B RSMTs represents the largest class of RSMTs, which can catalyze a wide range of reaction including phosphinates, unactivated sp^3-hybridized carbons, and sp^2-hybridized carbons. Van der Donk lab first proposed that methylation can be performed from methylcobalamin (MeCbl) to the substrate radical for these B_{12} dependent methyltransferases more than ten years ago.[23] However, the poor solubility upon overexpression and the substrate complexity have hampered the reconstitution and detailed *in vitro* studies of these enzymes for a long time. Recently, Booker's lab is able to develop a strategy to enhances the uptake of cobalamin into the cytoplasm and improves the solubility of the target enzymes significantly by co-overexpression of Cbl uptake protein.[24] Some of the earlier pioneer work in this area includes PhpK for the phosphinic acid P-methylation,[25] GenK and Fom3 for the alcohol C-methylation,[26-29] ThnK for the iterative C-methylation to form the ethyl group,[30] TsrM for the aromatic C-methylation,[31-32] PoyC for the valine residue C-methylation,[33-34] CysS for sequential methylations in cystobactamid biosynthesis (Figure 1.4). [35]

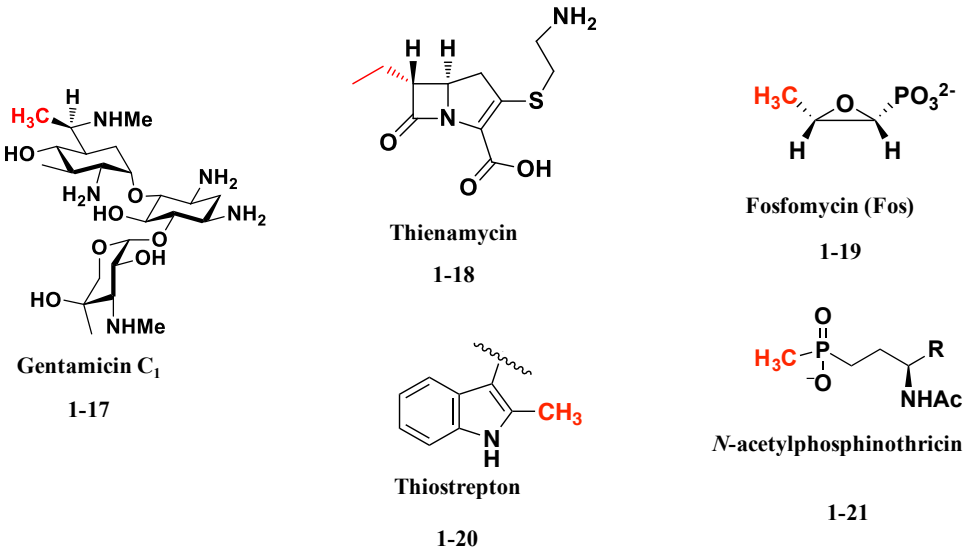


Figure 1.4 Selected secondary metabolites whose biosynthesis include class B RSMTs that have been reconstituted *in vitro*. Carbon atoms that are introduced by class B RSMTs are highlighted by red.

Class C RSMTs share high sequence similarity with HemN at the C-terminus in addition to the classical TIM barrel radical SAM domain. Class C RS can methylate sp^2-hybridized carbon centers, but do not have the two conserved cysteines like RlmN, Cfr in the sequence. In the crystal structure of HemN, there are two simultaneously bound SAM molecules in the active site.[36] Since, many residues in Class C RSMTs are conserved in HemN, including the residues for SAM binding, it has proposed that Class C RSMTs uses one molecular of SAM to generate 5′-dA radical and another molecule to provide methyl group. This suggests a possible mechanism in which the 5′dA• generated by the first SAM abstracts a hydrogen atom from the methyl group of the second SAM, which then can attack the unsaturated carbon to initiate the methylation.[20] Qi's group showed an alternative mechanism for NosN that SAM is first converted to 5′-

methylthioadenosine (MTA) and then was used as a direct methyl donor after it was abstracted a hydrogen atom by 5′-dA radical, which is different than what Booker's group observed.[37-38] Due to the poor accessibility of the complicated substrates and a lack of knowledge of the step along the biosynthetic pathway, there are few systems reconstituted for the Class C RSMTs. Recent studies also revealed that Class C RSMTs can not only methylate sp^2-hybridized carbon but also form cyclopropane rings or ester linkage (Figure 1.5). These include enzymes that catalyze the formation of cyclopropane rings formation (Jaw5 in jawsamycin biosynthesis, YtkT in yatakemycin biosynthesis, C10P and C10Q in CC-1065 biosynthesis),[39-41] methylation of sp^2-hybridized carbon (TbtI in thiomuracin biosynthesis, MenK/MqnK in menaquinone biosynthesis),[42-43] methylation and ester linkage (NosN in nosiheptide biosynthesis).[38]

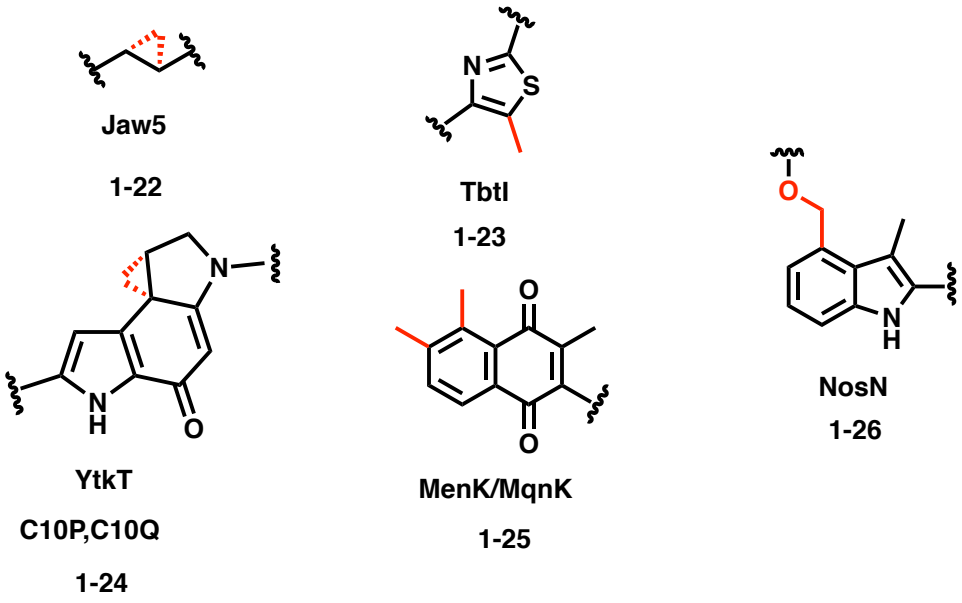


Figure 1.5 Class C radical SAM methyltransferases catalyzed modification (highlight in red) in the core structure of selected secondary metabolites.

2. RECONSTITUTION OF B_{12}-DEPENDENT RADICAL SAM METHYLTRANSFERASE CysS CATALYZED *t*-BUTYLATION BY ITERATIVE METHYLATION*

2.1. Introduction

Natural products with branched alkoxy groups play an important role in the development of bioactive compounds. In addition, the *t*-butyl group has fascinated organic chemists for more than a century and has played a major role in mechanistic studies on carbocation chemistry, organic substitution reactions, and the design and characterization of theoretically interesting molecules such as the remarkable tetra *t*-butyl tetrahedrane.[44] Although numerous *t*-butyl group substituted terpenes, polyketides and peptides have been identified, experimental studies on the biosynthesis of *t*-butyl groups are still at an early stage and many of the mechanistic proposals in the literature have not been adequately experimentally tested.[45]

For the ginkgolides and several other *t*-butyl substituted terpenes, the *t*-butyl group is formed by a double bond methylation using S-adenosylmethionine (Figure 2.1A).[46] Formation of the *t*-butyl group in the coumarin swietenone is proposed to involve carbocation insertion into a CH bond to give a cyclopropyl intermediate, which

---

* Reprinted in parts with permission from "Biosynthesis of Branched Alkoxy Groups: Iterative Methyl Group Alkylation by a Cobalamin-Dependent Radical SAM Enzyme" by Wang, Y.; Schnell, B.; Baumann, S.; Müller, R.; Begley, T. P. **2017**, *J. Am. Chem. Soc.* 139,1742-1745. Copyright 2017 American Chemical Society. Reprinted from Methods Enzymol, 606, Wang, Y.; Schnell, B.; Muller, R.; Begley, T. P., Iterative Methylations Resulting in the Biosynthesis of the *t*-Butyl Group Catalyzed by a B_{12}-Dependent Radical SAM Enzyme in Cystobactamid Biosynthesis, 199-216, Copyright (2018), with permission from Elsevier.

then undergoes acid mediated ring-opening (Figure 2.1B).[47] The biosynthesis of pivalic acid, a starter unit in the biosynthesis of *t*-butyl substituted polyketides, is mediated by a vitamin B_{12}-dependent mutase (Figure 2.1C).[48-49] Recently, a B_{12}/radical SAM mediated conversion of isopropyl glycine to *t*-butyl glycine in the polytheoamide propeptide was reported (Figure 2.1D).[33-34] A polyketide synthase is also found to catalyze coupled decarboxylation and methylation reactions to form pivaloyl-ACP from dimethylmalonyl-ACP in the apratoxin A biosynthesis(Figure 2.1E).[50] However, all of these enzymes catalyze a single methyl transfer to form the *t*-butyl group from an isopropyl group.

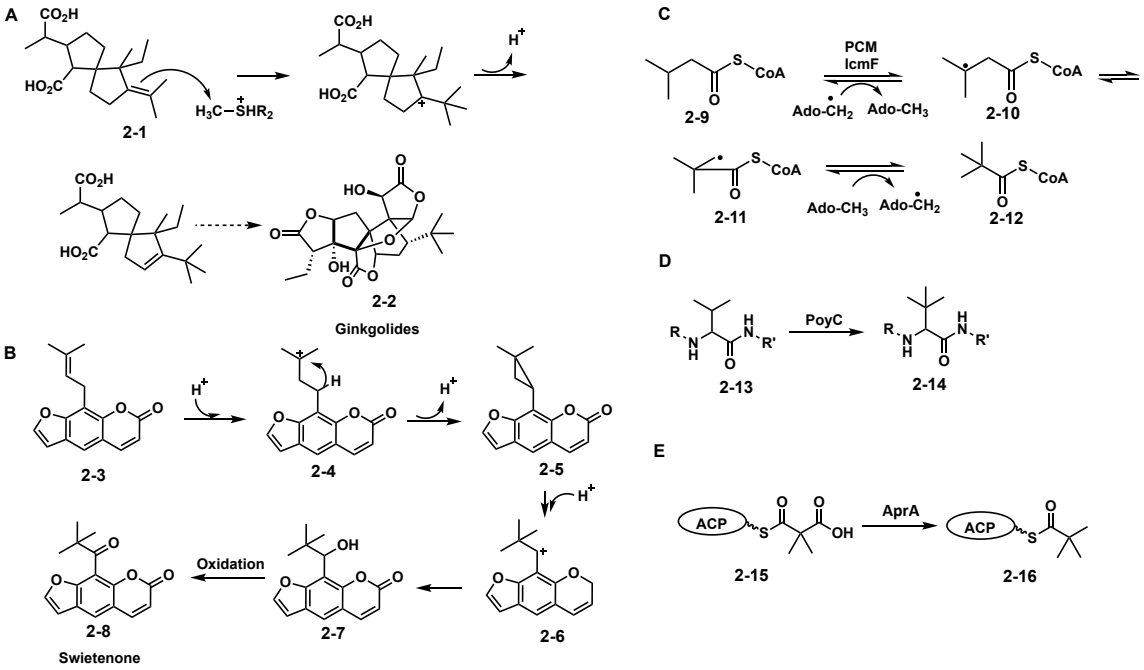


Figure 2.1 Mechanistic proposals for the formation of the *t*-butyl group in representative natural products.

The cystobactamids are a novel class of isopropyl substituted antibacterial compounds produced by myxobacteria.[51] The biosynthetic gene cluster has been identified and sequence analysis suggested that CysS is a cobalamin-dependent radical SAM methyltransferase, potentially involved in the iterative methylations of the 3-methoxy-4-aminobenzoic acid moieties of cystobactamid **2-19** (Figure 2.2). Some minor derivatives exhibit methyl, ethyl, isopropyl, and sec-butyl groups (Stephan Hüttel and R.M., unpublished results) supporting the hypothesis of CysS being an enzyme iteratively adding methyl-groups to its substrate.

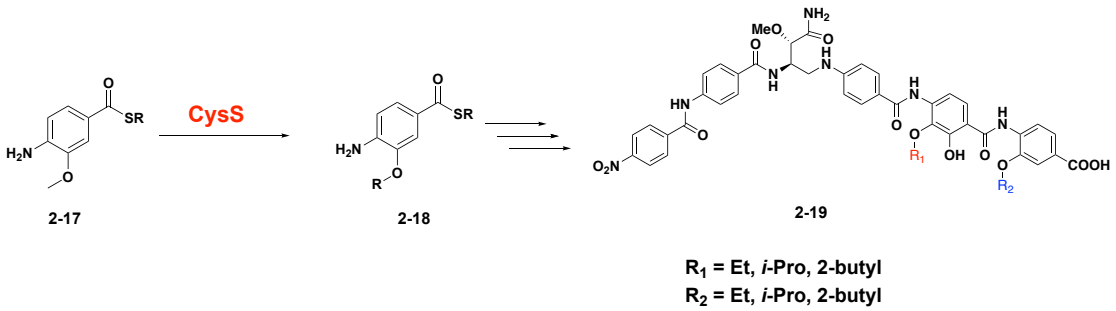


Figure 2.2 Proposed formation of branched alkoxy groups of cystobactamids by CysS-catalyzed iterative methylations of a methyl ether.

Here we successfully reconstitute CysS *in vitro* and demonstrate that this enzyme can assemble isopropyl, *sec*-butyl, and *t*-butyl groups by sequential methylations of a methyl group. To our knowledge, this is the first example of an isopropyl, *sec*-butyl, and a *t*-butyl group biosynthesis from a methyl group using radical chemistry.

2.2. Result and discussion

2.2.1. Reconstitution of CysS activity with minimal substrates

CysS was cloned into a pET28b vector and co-expressed with a plasmid encoding the *suf* operon ([4Fe-4S] biosynthesis)[52] in *Escherichia coli* BL21 (λDE3). The protein was then purified, under anaerobic conditions, by Ni-NTA affinity chromatography. The UV−visible spectrum of purified CysS revealed a 420 nm shoulder, typical of a bound Fe/S cluster (Figure 2.3). Iron and sulfide analysis yielded 2.5 irons and 2.8 sulfides per monomer of CysS, demonstrating partial cluster formation in the overexpressed protein.

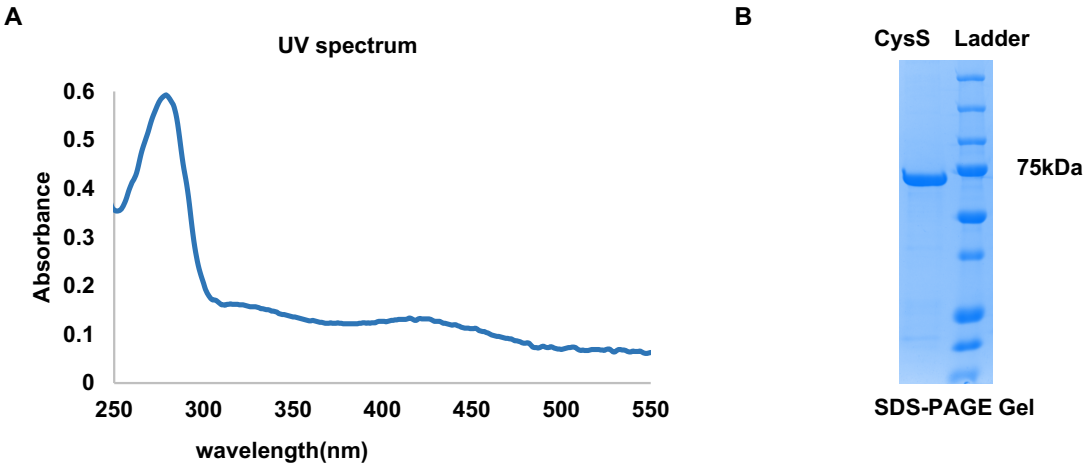


Figure 2.3 (A) UV spectrum of purified CysS (B) SDS/PAGE analysis of purified CysS.

Several *p*-aminobenzoic acid (PABA) analogs were tested as substrates for CysS. None gave the desired methylated product as indicated by LC-MS analysis. Further analysis of the cystobactamid biosynthesis cluster suggested the coenzyme A or the acyl

carrier protein thioester of **2-17** (CysG)[51] as possible CysS substrates. To test this proposal, N-acetylcysteamine thioester **2-20** was synthesized and incubated with CysS, SAM, MeCbl, and flavodoxin/flavodoxin reductase/NADPH (Figure 2.4). LC-MS analysis of the resulting reaction mixture demonstrated the formation of the ethyl ether **2-21**(Figure 2.5A). This was further confirmed by co-elution of the reaction product with a synthesized sample of **2-21** (Figure 2.5D). When the ethyl ether **2-21** was incubated with CysS, the isopropyl ether **2-22** was detected by LC-MS analysis (Figure 2.5C).

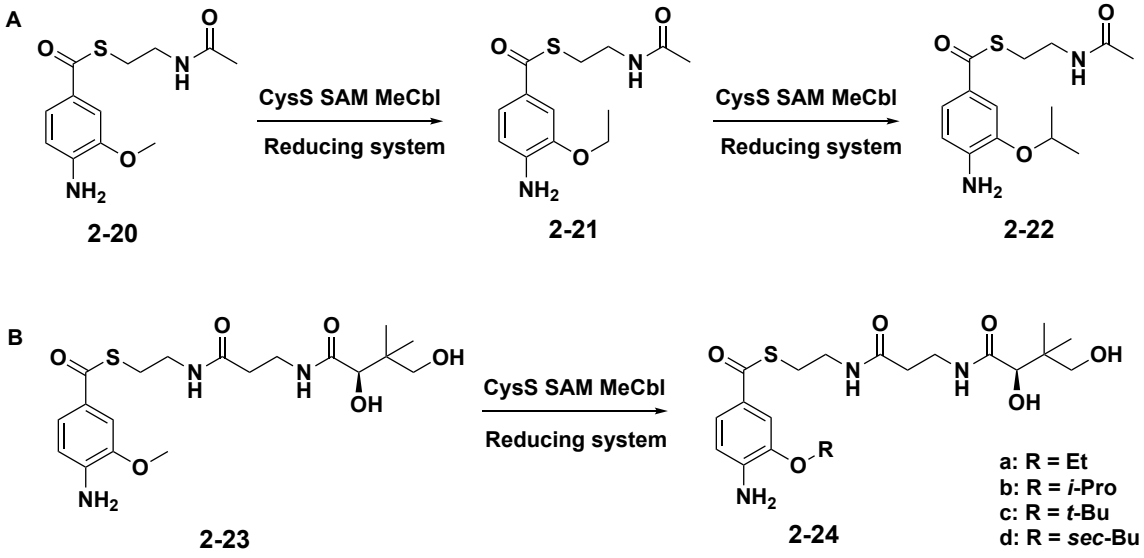


Figure 2.4 Identification of two substrates for CysS. A competition reaction with a 1:1 mixture demonstrated that **2-23** is 47 times more reactive than **2-20**.

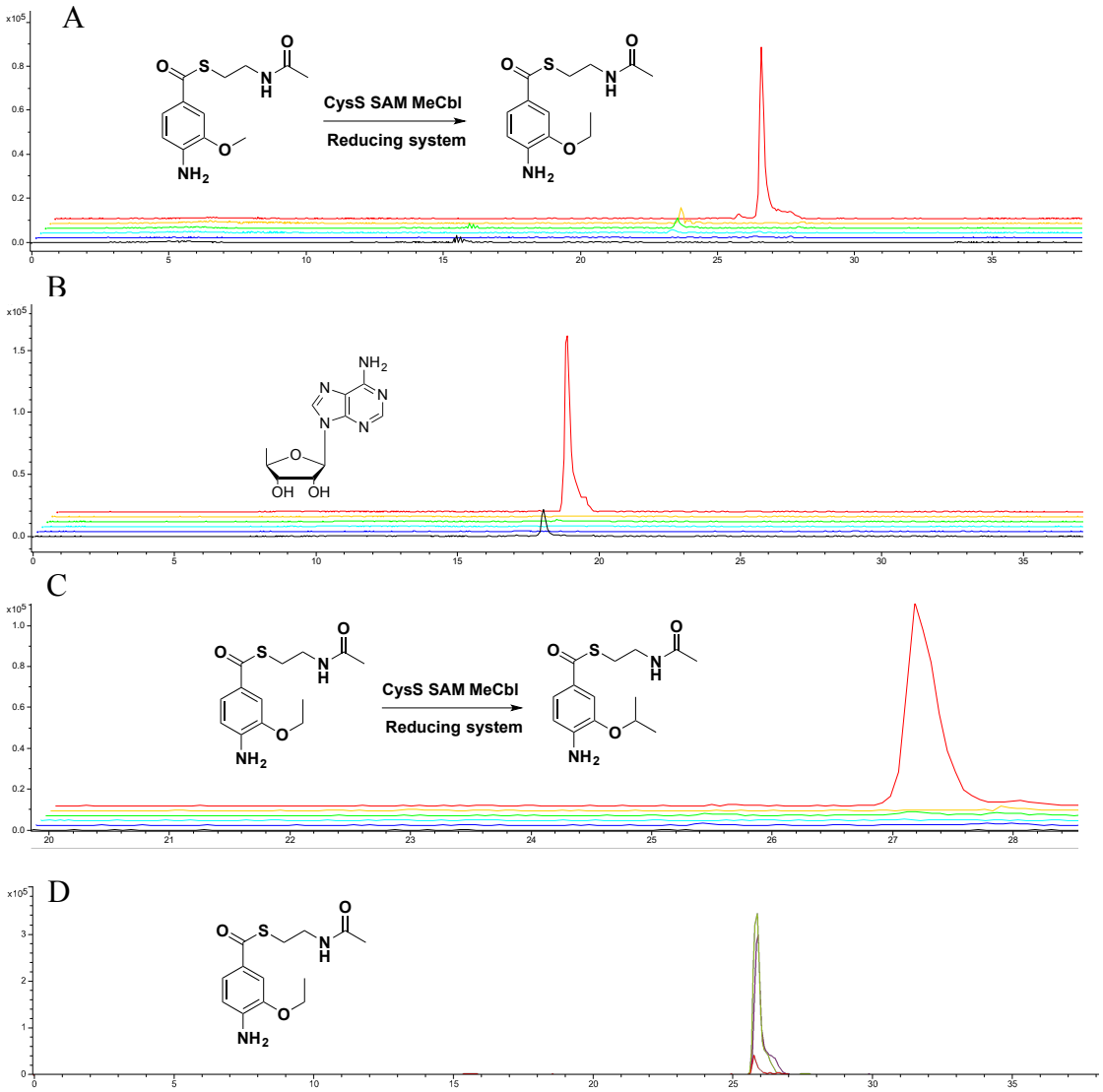


Figure 2.5 LC-MS detection of the CysS reaction products using methyl ether **2-21** and ethyl ether **2-22**. Red trace is the full reaction where all the components are present. Orange, cyan, blue, black and green traces respectively are for reaction mixtures where either CysS, MeCbl, SAM, substrate, or reducing system flavodoxin/flavodoxin reductase/NADPH) is absent. (A) EIC of compound **2-21** $[M+H]^+$ (283.11±0.02) (B) EIC of 5′-dA $[M+H]^+$ (252.10±0.02) (C) EIC of compound **2-22** $[M+H]^+$ (297.13±0.02) (D) EIC of compound **2-22** co-elution with an authentic sample $[M+H]^+$ (283.11±0.02). Red trace is the full enzymatic reaction. Purple trace is compound standard **2-22**. Green trace is the co-elution of the enzymatic product and the synthetic standard.

Pantetheinyl thioester **2-23** was a better substrate for CysS and iterative methylations to give the ethyl, isopropyl, and the butyl ethers were detected by LC-MS analysis (Figure 2.6). Small amounts of the ethers **2-24a** and **2-24b** were detected in the absence of the reducing agent suggesting that some of the purified enzyme contained the reduced [4Fe−4S] cluster. To confirm the structures of **2-24a-c**, authentic samples of these compounds were synthesized. The enzymatic products matched the synthetic standards in terms of retention time, exact mass, and fragmentation pattern (Figure 2.12, 2.13). In addition, CysS catalyzed the conversion of synthetic **2-24a** to **2-24b-d** and the conversion of synthetic **2-24b** to **2-24c, d** (Figure 2.15). The second component in the extracted ion chromatogram for the *t*-butyl ether **2-24c** (Figure 2.6C, D) was identified as the sec-butyl ether **2-24d** by co-migration with an authentic standard of **2-24d** (Figure 2.16).

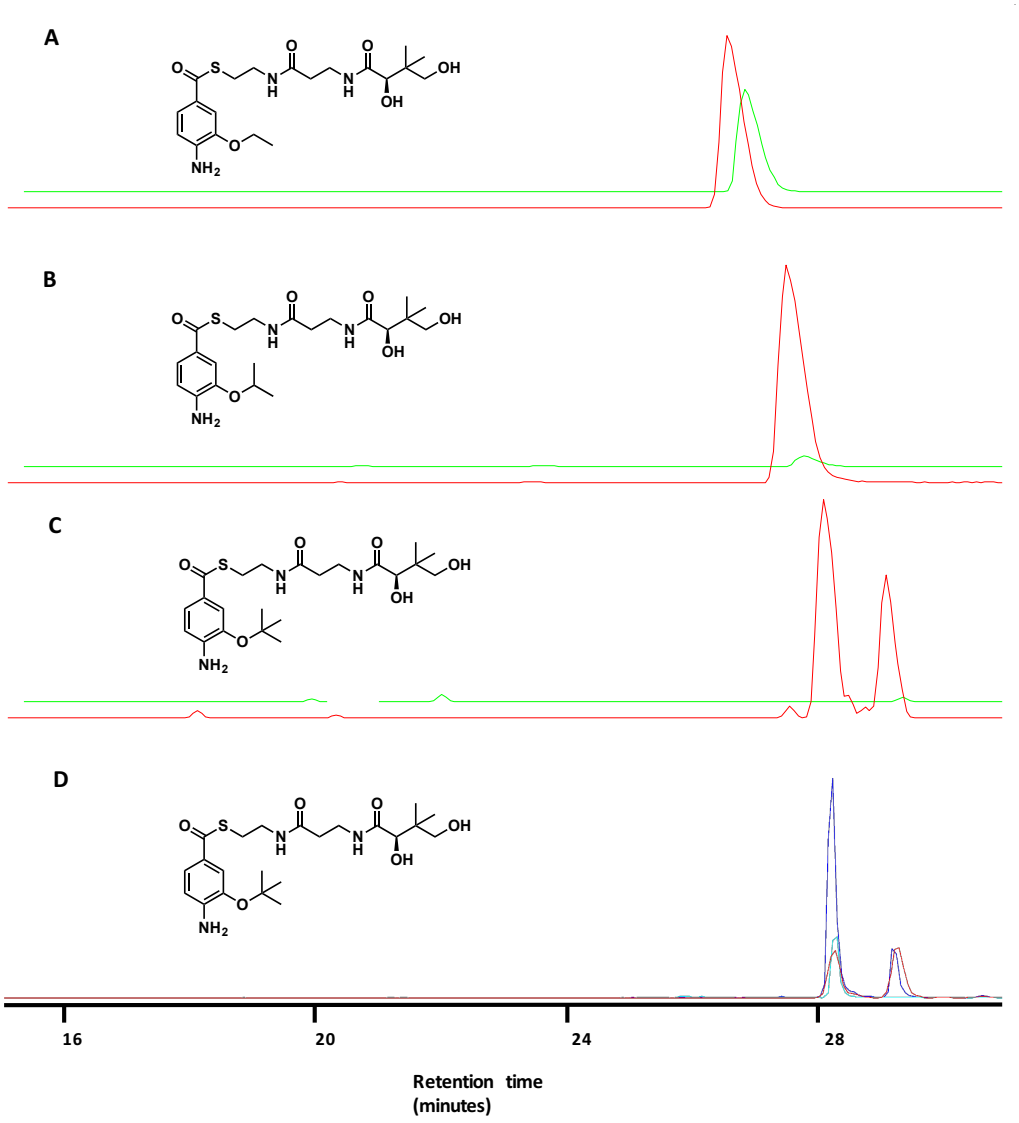


Figure 2.6 LC-MS analysis of CysS-catalyzed iterative methylations of methyl ether **2-23**. Red trace is for the complete reaction mixture. Green trace is for reaction mixtures where the reducing system (flavodoxin/ flavodoxin reductase/NADPH) is absent. Ethyl ether **2-24a**, isopropyl ether **2-24b** were not formed in the control reactions lacking CysS, SAM, or MeCbl. (A) Extracted ion chromatograms (EICs) of the ethyl ether **2-24a** $[M + H]^+$ (442.20 ± 0.02). (B) EICs of the isopropyl ether **2-24b** $[M + H]^+$ (456.22 ± 0.02). (C) EICs of the *t*-butyl ether **2-24c** $[M + H]^+$ (470.23 ± 0.02). (D) EICs of $[M + H]^+$ (470.23 ± 0.02) showing co-migration with a synthesized sample of **2-24c**. Cyan trace is the *t*-butyl ether standard. Blue trace is co-elution of the enzymatic product and synthetic standard. The second component in the extracted ion chromatogram for the *t*-butyl ether **2-24c** (panels C, D) was identified as the sec-butyl ether **2-24d**.

Various [4Fe−4S] cluster reducing agents were tested in addition to the flavodoxin/flavodoxin reductase/NADPH. NADPH/methyl viologen, a commonly used electron source for cobalamin-dependent radical SAM enzymes, gave similar activity. However, dithionite or the combination of methyl viologen and dithionite gave a significantly lower activity.[53] Buffer thiols inactivate the substrate by trans thioesterification and need to be avoided.

Quantitative analysis of the enzymatic reaction mixture (CysS, methyl ether **2-23**, flavodoxin/flavodoxin reductase/NADPH, reaction run to completion) by LC-MS showed that 1 equiv of enzyme undergoes >2 turnovers, generating around 2.0 equiv of 5′-dA, 2.0 equiv of SAH, 1.4 equiv of ethyl ether **2-24a** and 0.3 equiv of isopropyl ether **2-24b** (Figures 2.6 and 2.7). The *t*-butyl ether **2-24c** was detected only when the concentration of the isopropyl ether **2-24b** was >23 μM. The ratio of 5′-dA to SAH was close to 1, suggesting that two molecules SAM were consumed for each methylation reaction and that the uncoupled production of 5′-dA is low. This is consistent with SAM functioning as the source of both the adenosyl radical and the methyl group and was further supported by LC-MS analysis of a reaction mixture containing CD_3-SAM which demonstrated CD_3 incorporation into the ethyl ether **2-24a** and the isopropyl ether **2-24b** (Figure 2.8).

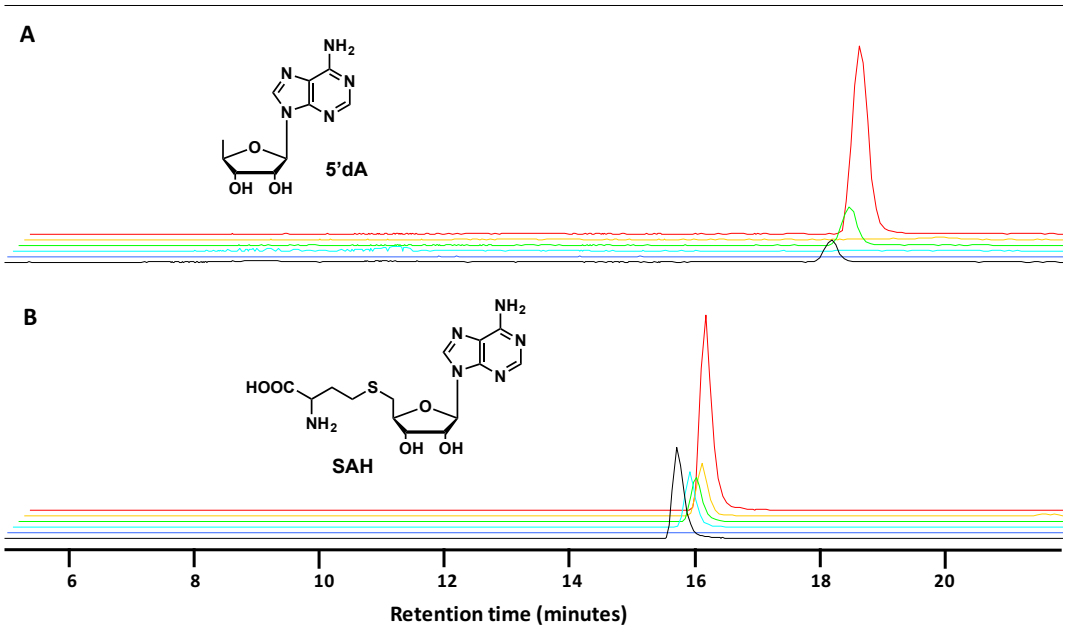


Figure 2.7 LC-MS detection of 5′-dA and SAH in the CysS-catalyzed iterative methylations of the methyl ether **2-23**. Red trace is for the complete reaction mixture. Green trace is for reaction mixtures where the reducing system (flavodoxin/flavodoxin reductase/NADPH) is absent. (A) EICs of 5′-dA $[M + H]^+$ (252.11 ± 0.02). (B) EICs of SAH $[M + H]^+$ (385.13 ± 0.02). Yellow, green, cyan, and black traces are for reaction mixtures where either CysS, reducing system, MeCbl or substrate is absent. The product ratio was determined by calibrating the signal intensity with known concentrations of standards

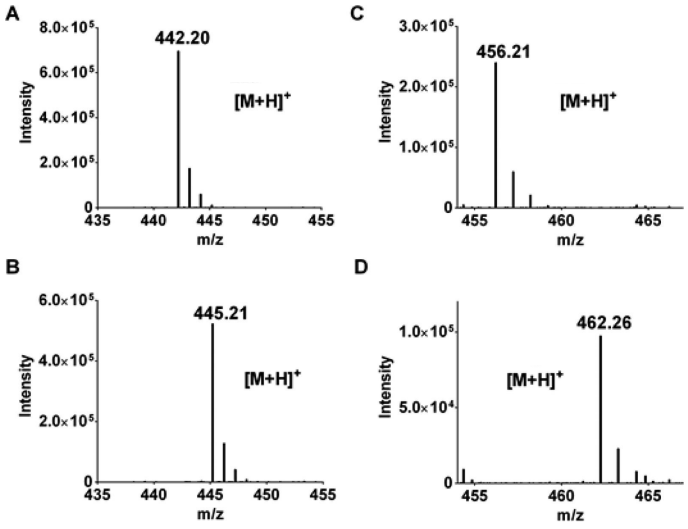


Figure 2.8 MS analysis of a reaction mixture in which CH_3-SAM is replaced with CD_3-SAM showing CD_3 incorporation into the ethyl and isopropyl ethers of **2-24a** and **2-24b**, respectively. Panels A and C: Mass spectra of **2-24a** and **2-24b** formed from CH_3-SAM. Panels B and D: Mass spectra of 22a and 22b formed from CD_3-SAM.

A mechanistic proposal for the CysS-catalyzed reaction, based on the proposed mechanisms for GenK and ThnK is shown in Figure 2.9. After initial formation of methylcobalamin by SAM mediated methylation, reductive cleavage of SAM by the $[4Fe-4S]^{+1}$ cluster generates the 5′-deoxyadenosyl radical. This abstracts a hydrogen atom from the methyl group of the substrate **2-23** to give radical **2-25**, which then undergoes a radical substitution with methyl cobalamin to give the ethyl ether **2-24a**. An analogous methyl transfer, by a radical substitution mechanism, has precedence in cobalamin model chemistry.[54] Regeneration of MeCbl from Cbl(II) can be achieved by reduction to Cbl(I) by the $[4Fe-4S]^{+1}$ cluster followed by SAM-mediated methylation. Repetition of this sequence results in the successive formation of the isopropyl, *t*-butyl, and *sec*-butyl ethers of **2-23**.

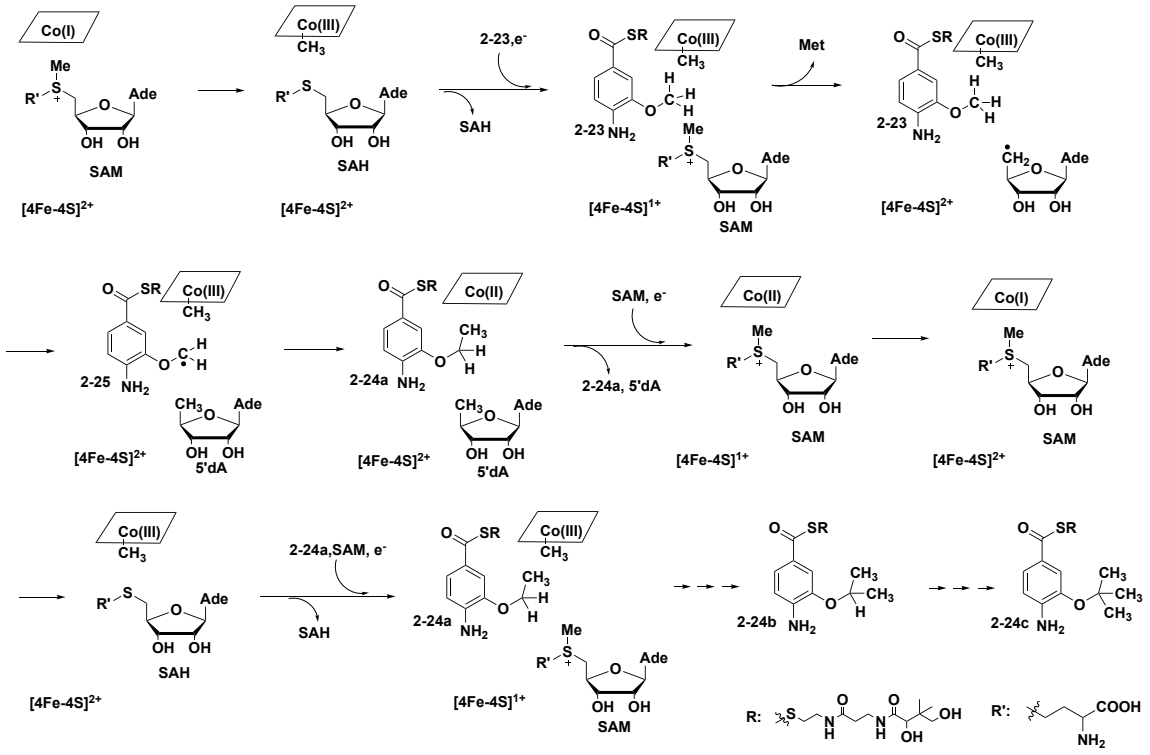


Figure 2.9 Proposal for CysS-catalyzed iterative methylations to form branched alkoxy groups. The mechanism assumes two different SAM binding sites. It is also possible that SAM binds to a single site and that the position of the sulfonium moiety is altered by a protein conformational change.

2.2.2. Reconstitution of CysS activity with native substrates

The most logical substrate for CysS is the 4-amino-3-methoxybenzoate thioester with the peptidyl carrier protein of the polypeptide synthase.[51] This was synthesized as shown in Figure 2.10. Condensation of **2-26** with coenzyme A yielded the CoA ester **2-27**. The phosphopantetheinyl thioester moiety of **2-27** was then transferred to CysG1 in a reaction catalyzed by phosphopantetheinyl transferase.

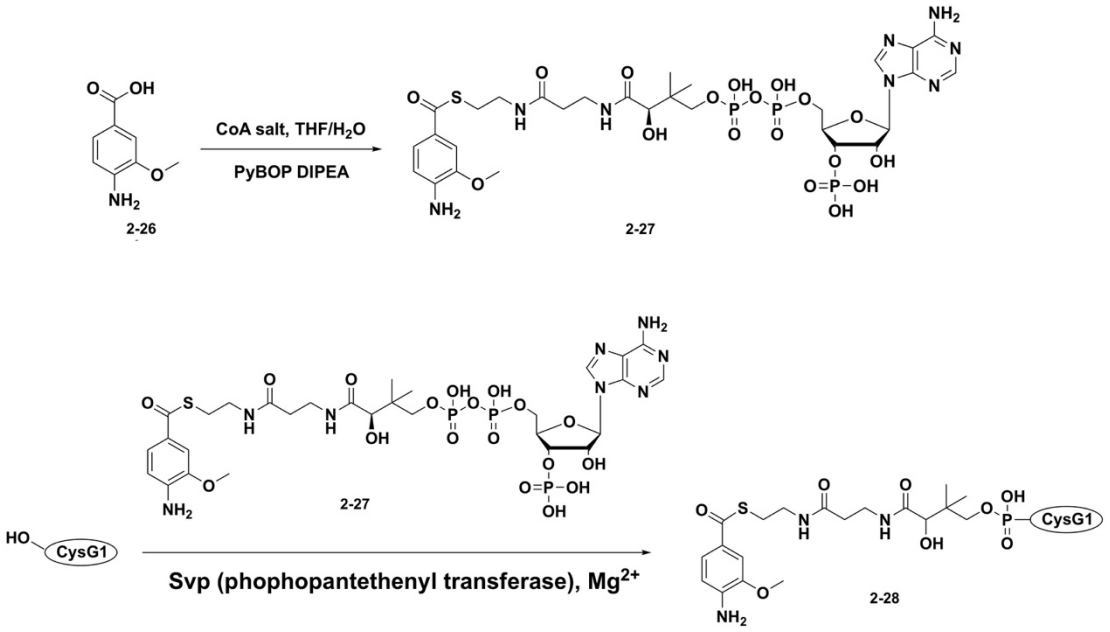


Figure 2.10 Synthesis of the 4-amino-3-methoxybenzoate (AMB) CysG1 thioester **2-28**. Adapted with permission from 55.

When AMB-CysG1 thioester was incubated with CysS, MeCbl, SAM and methyl viologen, dimethylated **2-29a** and trimethylated products **2-29b** were detected by LC-MS (the formation of trimethylated products required higher enzyme concentrations) (Figure 2.11). This further demonstrated that the peptide carrier protein thioester of **2-28** is the native substrate of CysS.[55]

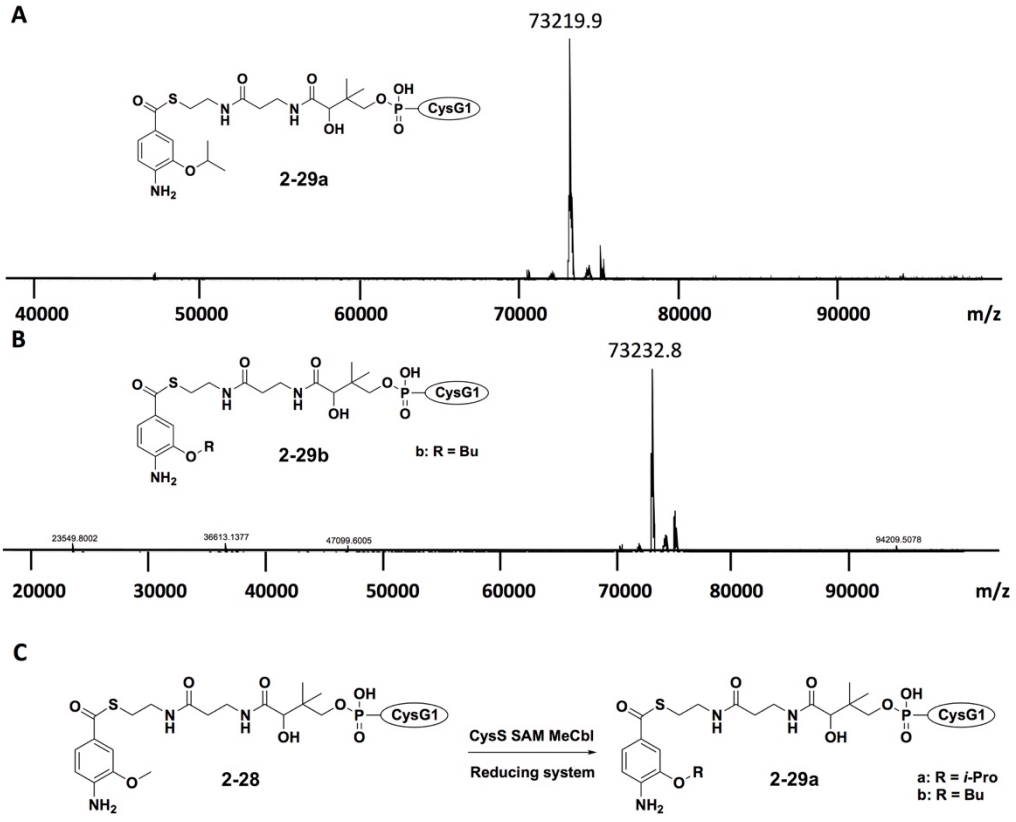


Figure 2.11 The CysS-catalyzed reaction of the AMB-CysG1 thioester. (A) MS of the dimethylated product (APB-CysG1 thioester **2-29a**, calculated monoisotopic mass: 73217.8Da, experimental mass: 73218.0Da), (B) MS of the trimethylated product (ABB-CysG1 thioester **2-29b**, calculated monoisotopic mass: 73231.7Da, experimental mass: 73232.7Da), (C) The reaction catalyzed by CysS. Adapted with permission from 55.

2.3. Conclusion

Here we describe the successful in *vitro* reconstitution of CysS and demonstrate that this enzyme can assemble isopropyl, *sec*-butyl, and *t*-butyl groups by sequential methylations of a methyl group. To our knowledge, this is the first example of an isopropyl, *sec*-butyl, and a *t*-butyl group biosynthesis from a methyl group using radical chemistry. This biosynthetic strategy in principle enables the host myxobacterium to

biosynthesize a combinatorial antibiotic library of 25 cystobactamid analogs. Analogous methyl-based combinatorial libraries could potentially be generated in natural products such as kedarcidin and quinomycin SW-163s, which have ethyl, isopropyl, and *sec*-butyl substituents on a thiomethyl group.[56-57]

2.4. Experimental methods

2.4.1. overexpression and purification of CysS

A 15mL overnight culture was grown in LB medium in the presence of 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. This was then added to 1.5 L of LB medium containing 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. The cultures were incubated at 37°C with shaking (180 rpm) until the OD600 reached 0.45. The culture was then incubated at 4°C without shaking for 1 hour. Expression was induced by the addition of 100 μM IPTG followed by ferrous ammonium sulfate (120 mg) and cysteine (120 mg) and the culture was grown at 15 °C with shaking (110 rpm) overnight. The cells were harvested by centrifugation and stored in liquid nitrogen until further use. All steps for protein purification were carried out in an anaerobic chamber (COY laboratories). Cell pellets (~12 g) were thawed and resuspended in 30 mL lysis buffer (100 mM Tris-HCl, pH 7.5) in the presence of lysozyme (0.2 mg/mL) and benzonase (100 units). This mixture was cooled in an ice-bath with continuous stirring for 1-2 hrs. The suspension was then sonicated to complete cell lysis for 5 × 120 s using a Misonix Sonicator XL-2000 (setting = 9). Cell debris was removed by centrifugation, and the lysate was loaded onto a Histrap column pre-equilibrated with lysis buffer. The column was washed with 5-10 column volumes of wash buffer (100 mM Tris-HCl, 30 mM

Imidazole, pH 7.5). The CysS protein was then eluted using elution buffer (100 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH 7.5). The brown-colored fractions, containing the desired protein, were pooled and buffer exchanged to final buffer (100 mM Tris-HCl, pH 7.5 at 25 °C, 30% glycerol) using Econo-Pac 10DG size exclusion chromatography. The purified CysS was divided into aliquots in eppendorf tubes, frozen in liquid nitrogen, and stored at −80 °C, submerged in liquid nitrogen, until future use. Each tube has a pin hole in the cap to allow for gas release during thawing. Protein concentration was measured by the absorbance at 280 nm (A280) when the protein was exposed to air or the [4Fe-4S]2+ cluster was reduced to [4Fe-4S]1+ by dithionite (prevent the interference from absorbance from [4Fe-4S]2+) with an extinction coefficient calculated using the ProtParam tool of the ExPASy proteomics server. Typical yields were approx. 5 mg per liter.

2.4.2. Iron and sulfide quantification of CysS

Iron content was analyzed using the ferene assay (Cooper et al., 2013). The following reagents were prepared: A, 1.35 g SDS, 30 mL water, 450 µL saturated sodium acetate; B, 540 mg L-ascorbic acid, 18 mg sodium metabisulfite, 11.2 mL water, 800 µL saturated sodium acetate; C, 18 mg ferene in 1 mL water. Guanidine hydrochloride (8M, 300 µL), Reagent A (300 µL), and Reagent B (300 µL) were added to 300 µL of a CysS solution. After mixing by inversion, the samples were incubated for 15 min at 37 °C and then 15 µL of Reagent C was added. The samples were incubated for 5 min at 25 °C and centrifuged for 5 min at 25 °C. The absorbance at 593 nm of the supernatant was recorded using a Varian Cary Bio 300 UV-Visible Spectrophotometer.

The iron content was finally determined by comparing the reading to a standard curve that was generated under identical conditions using FeCl3 in 2% HNO3 with a concentration range from 0 to 200 μM.

Sulfide content was determined using the methylene blue assay.[6] To 300 μL of the assay solution containing CysS, 1 ml of 1% (w/v) zinc acetate was added followed by 50 μL of 3 M NaOH. This mixture was agitated gently and 250 μL of 0.1 % N, N-dimethyl-p-phenylenediamine (DMPD) monohydrochloride in 5 M HCl and 50 μL of 23 mM FeCl3 in 1.2 M HCl were added. The resulting solution was mixed vigorously for 30 mins. The samples were then centrifuged for 10 min at 25 °C. The supernatant was collected and the absorbance at 670 nm was recorded using a Varian Cary Bio 300 UV-Visible Spectrophotometer. The sulfide content was determined by comparing this reading to a standard curve that was generated under identical conditions using a fresh solution of sodium sulfide (Na2S) in 0.1 M NaOH with a concentration range from 0 to 100 μM.

2.4.3. *In vitro* reconstitution assay with CysS:

All CysS enzymatic reactions were carried out in an anaerobic chamber containing 95% nitrogen and 5% hydrogen. A typical enzymatic reaction was performed in 100 mM phosphate buffer, pH 7.5 containing 55 μM of CysS with 0.75 mM MeCbl, 2.5 mM SAM, 0.9 mM substrate with different reducing systems, flavodoxin, flavodoxin reductase system (25 μM FldA, FPR; 50 μM FAD, FMN; 3 mM NADPH); 1 mM methyl viologen, 4 mM NADPH; 10 mM dithionite, 10 mM dithiothreitol (DTT); 1 mM methyl viologen, 4 mM NADPH, 50 mM DTT; 1 mM methyl viologen, 4 mM NADPH,

50 mM DTT. The reaction was incubated at room temperature for 12-15 hours. For LC-MS analysis the enzyme was removed by ultrafiltration using a 10kDa cut-off filter (VWR).

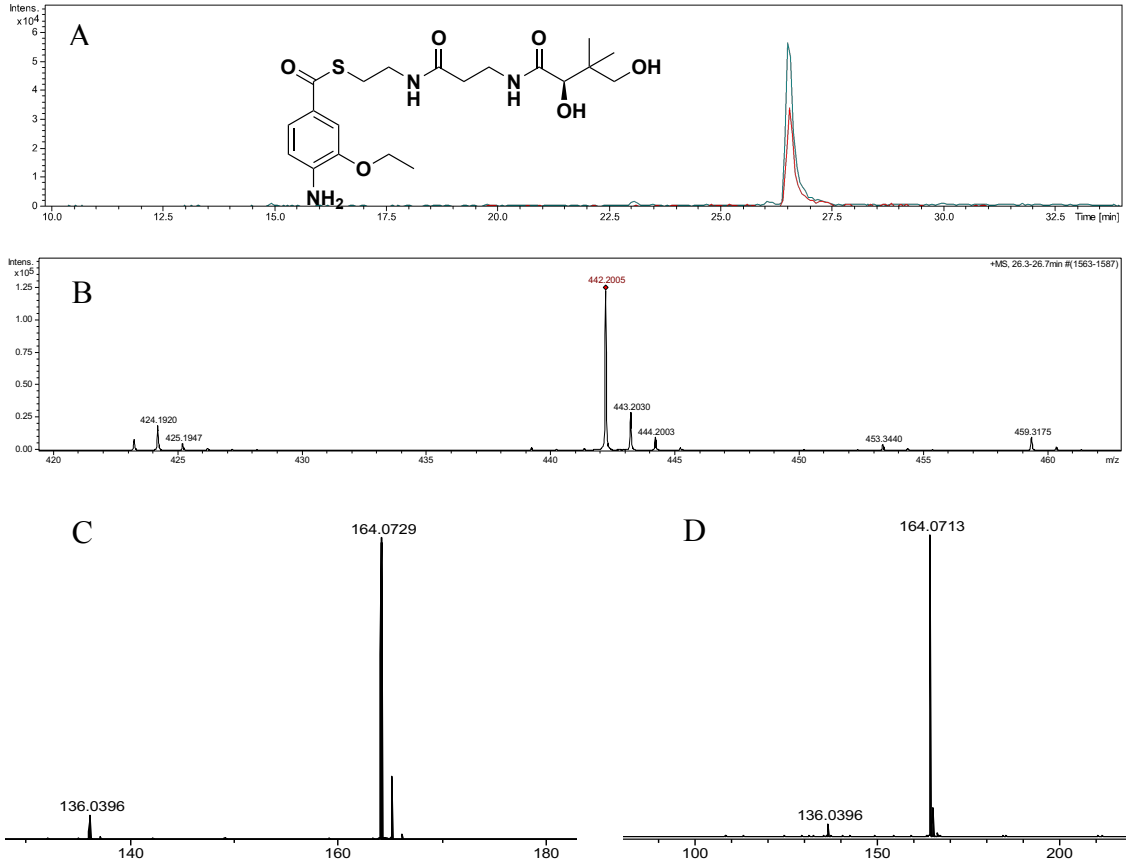


Figure 2.12 LC-MS detection of **2-24a** in the CysS reaction using methyl ether **2-23** as substrate. Red trace is the enzymatic full reaction. Green trace is the co-elution of the enzymatic product and synthetic standard of the ethyl ether **2-24a** (A) EIC of ethyl ether **2-24a** $[M+H]^+$ (442.20±0.02). (B) Mass spectrum of ethyl ether **2-24a** (exact mass: 442.2021, Accuracy: <1 ppm). (C) Mass fragments of the ethyl ether **2-24a** produced in the enzymatic reaction. (D) Mass fragments of synthetic standard of the ethyl ether **2-24a**.

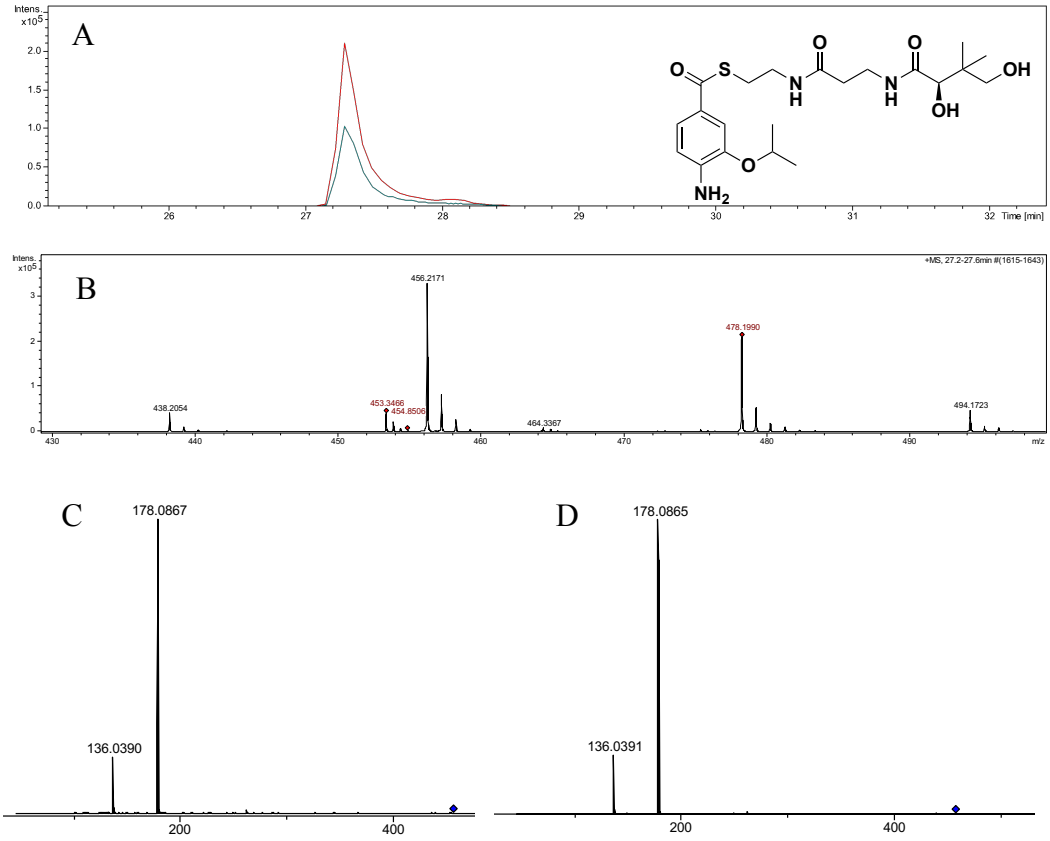


Figure 2.13 LC-MS detection of **2-24b** in the CysS reaction using methyl ether **2-23** as substrate. Red trace is enzymatic full reaction. Blue trace is synthetic standard of isopropyl ether **2-24b** (A) EIC of isopropyl ether **2-24b** $[M+H]^+$ (456.21±0.02). (B) Mass spectrum of isopropyl ether **2-24b** (exact mass: 456.2171, Accuracy: 1.7 ppm). (C) Mass fragments of the isopropyl ether **2-24b** produced in the enzymatic reaction. (D) Mass fragments of synthetic standard of isopropyl ether **2-24b**.

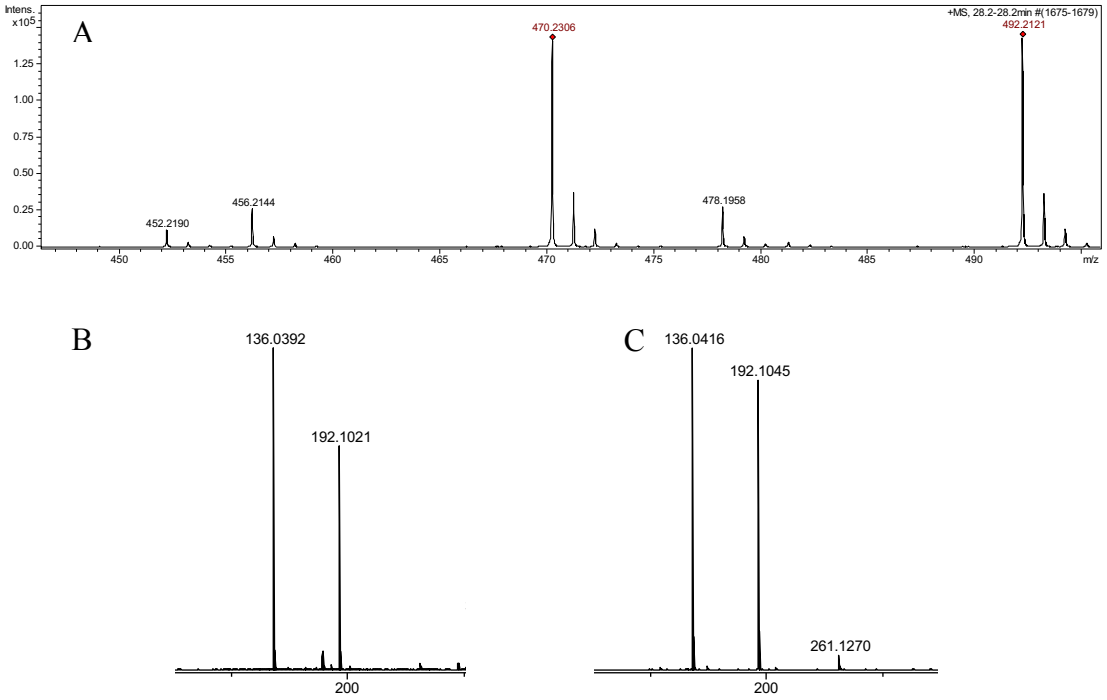


Figure 2.14 LC-MS detection of **2-24c** in the CysS reaction using methyl ether **2-23** as substrate. (A) Mass spectrum of *t*-butyl ether **2-24c** (exact mass: 470.2322, Accuracy: 3.4 ppm). (B) Mass fragments of *t*-butyl ether **2-24c** produced in the enzymatic reaction. (C) Mass fragments of synthetic standard of *t*-butyl ether **2-24c**.

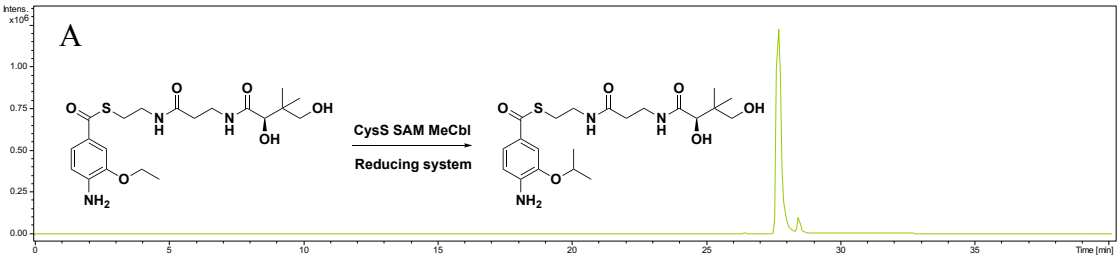


Figure 2.15 LC-MS detection of the CysS reaction products using ethyl ether **2-24a** (A), (B) or isopropyl ether **2-24b** (C). (A) EIC of isopropyl ether **2-24b** $[M+H]^+$ (456.21±0.02). (B), (C) EIC of butyl ethers **2-24c, 2-24d** $[M+H]^+$ (470.23±0.02).

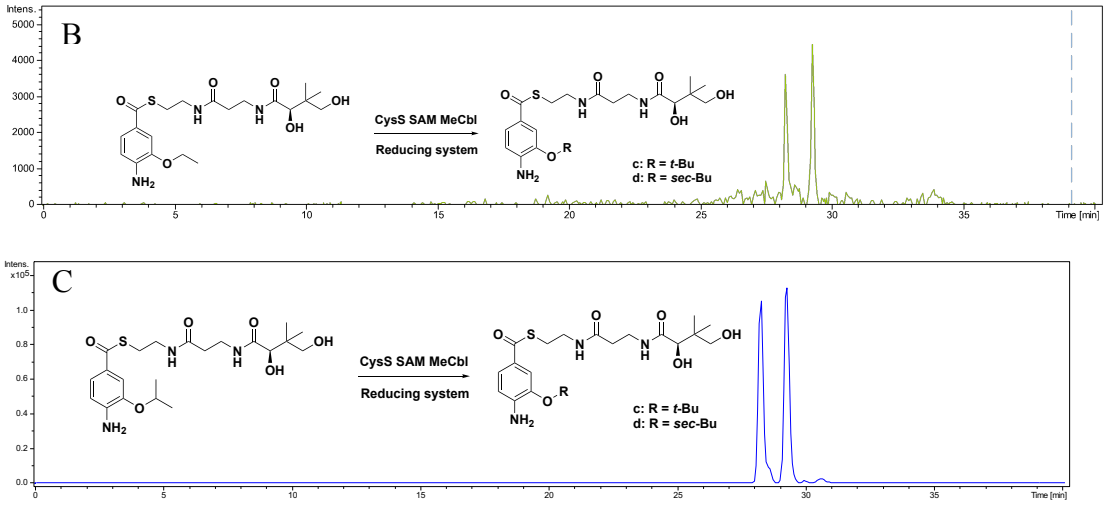


Figure 2.15 Continued

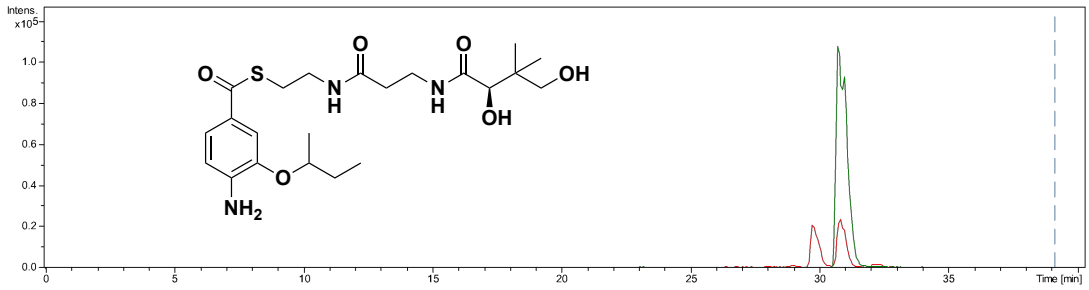


Figure 2.16 LC-MS detection of the CysS reaction products using isopropyl ether **2-24b** as substrate. The second peak was co-migrated with an authentic standard of *sec*-butyl ether **2-24d**. Red trace is EIC of mass (470.23±0.02) for enzymatic full reaction. Green trace is synthetic standard of *sec*-butyl ether **2-24d**.

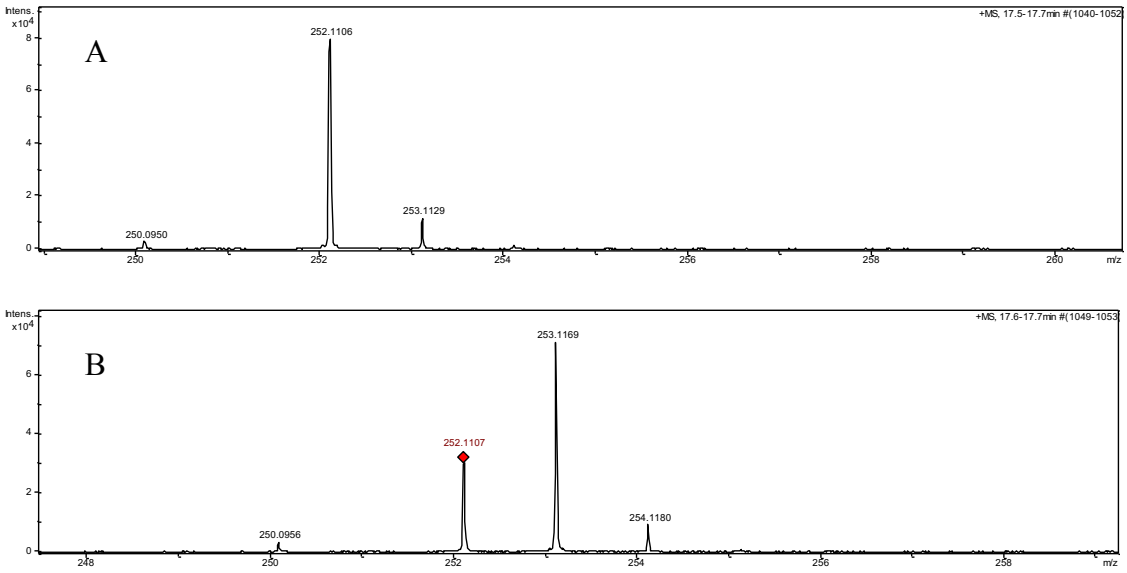


Figure 2.17 LC-MS detection of 5′-dA produced in the CysS reaction products using (A) methyl ether **2-23** or (B) CD_3-methyl ether **CD_3-2-23** as substrate.

2.4.4. LC-MS parameters for the analysis of the CysS reaction mixture

LC-ESI-TOF-MS was performed using an Agilent 1260 HPLC system equipped with a binary pump and a 1200 series diode array detector followed by a MicroToF-Q II mass spectrometer (Bruker Daltonics) using an ESI source in positive mode. Analysis was performed on a LC-18-T column (15 cm x 3 mm, 3 μm particles, Supelco).

LC conditions:

A- 5 mM Ammonium acetate buffer, pH 6.6

B- 75 % Methanol and 25 % Water.

LC method:

0min – 100% A, 2 min – 100% A, 4 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

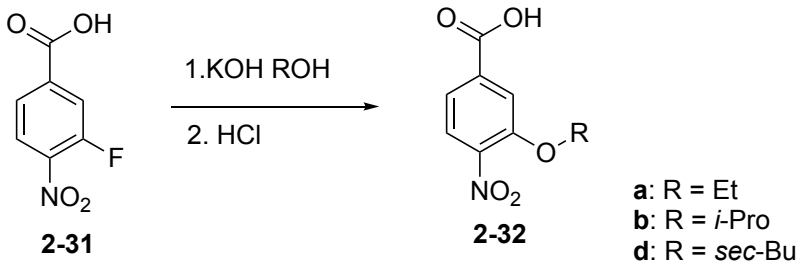
The mass spectrometer parameters for small molecules: Capillary, 4500 V; end plate offset, 500 V; nebulizer gas, 3.0 bar; dry gas, 10.0 L/min; dry gas temperature, 200°C; funnel 1 RF, 250.0 Vpp; funnel 2 RF, 300.0 Vpp; isCID, 0.0 eV; hexapole RF, 200 Vpp; quadrupole, Ion energy, 3.0 eV; low mass, 150 m/z; collision cell, collision energy, 8.0 eV; collision RF, 150.0 Vpp; transfer time, 80.0 μs; prepulse storage, 5.0 μs. Data were processed with DataAnalysis ver. 4.0 (Bruker Daltonics).

2.4.5. Quantification of enzymatic products in the CysS reaction mixture

The CysS assay, using methyl ether **2-23** was carried out under the above conditions with flavodoxin, flavodoxin reductase or methyl viologen, NADPH as reducing systems. Calibration curves were constructed by measuring the peak intensity of 5′-dA, SAH, synthetic ether **2-21**, **2-24a** and **2-24b** with known concentration using LC-MS. Then the calibration curves were used to measure the concentration of 5′-dA, SAH, **2-21**, **2-24a** and **2-24b** in the enzymatic reaction. Control reactions without CysS were used to assess the quantity of non-enzymatically produced SAH and this amount was subtracted out of the enzymatic reaction to give the final result. With our current enzyme preparation, 55 μM of CysS with 0.75 mM MeCbl, 2.5 mM SAM, 0.9 mM substrate and flavodoxin, flavodoxin reductase system (25 μM FldA, FPR; 50 μM FAD, FMN; 3 mM NADPH) produced 111 μM of 5′-dA, 109 μM of SAH, 78 μM of ethyl ether **2-24a**, 15 μM of isopropyl ether **2-24b**. The relative rate of **2-20** and **2-23** was measured by the amount of product formations of **2-21** and **2-24** by CysS with 1:1 ratio of **2-20** to **2-23**. 11 μM CysS was incubated with 0.3 mM MeCbl, 0.5 mM SAM, 0.5

mM 18, 0.5 mM 21, and 1 mM methyl viologen, 3 mM NADPH produced 0.4 uM of **2-21**, 15.8 μM of **2-24a**, 3.2 μM of **2-24b**.

2.4.6. General procedure for the synthesis of the 4-nitrobenzoic acids (2-32a, b, d)



2-31 **2-32**

a: R = Et
b: R = *i*-Pro
d: R = *sec*-Bu

Potassium hydroxide (5.0 mmol) was added to a solution of 3-fluoro-4-nitrobenzoic acid (2.0 mmol) 31 in excess of alcohol (10 ml) and the reaction mixture was slowly heating to reflux for several hours and monitored by TLC (Hexane and EtOAc solvent system). Then it was acidified using 2N HCl and extracted with ethyl acetate. The extracts were combined, washed with water, dried with MgSO4 and evaporated to give 4-nitro-benzoic acids **2-32a, b** and **d**.

3-ethoxy-4-nitrobenzoic acid (2-32a)

Compound **2-32a** was prepared by the above procedure using ethanol in 63% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.80 (m, 2H), 7.68 (d, J = 9.8 Hz, 1H), 4.26 (q, J = 7.0 Hz, 2H), 1.43 (t, J = 7.0 Hz, 3H). ^{13}C NMR (101 MHz, MeOD) δ 152.59, 136.72, 125.79, 122.58, 116.71, 106.43, 66.77, 14.72.

3-isopropoxy-4-nitrobenzoic acid (2-32b)

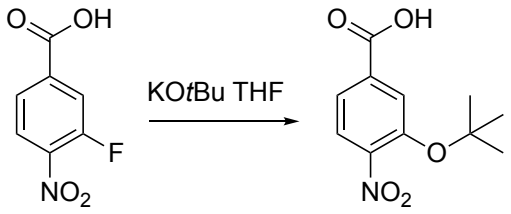
Compound **2-32b** was prepared by the above procedure using isopropyl alcohol in 90% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.81 (s, 1H), 7.76 (d, J = 8.3 Hz, 1H), 7.66

(d, J = 8.3 Hz, 1H), 4.80 (m, 1H) 1.36 (d, J = 6.0 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 172.96, 151.57, 136.51, 125.75, 122.57, 118.13, 74.13, 22.03, 20.85.

3-(*sec*-butoxy)-4-nitrobenzoic acid (2-32d)

Compound **2-32d** was prepared by the above procedure using 2-butanol in 37% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.87 – 7.72 (m, 2H), 7.66 (d, J = 8.3 Hz, 1H), 4.68 – 4.51 (m, 1H), 1.82 – 1.60 (m, 2H), 1.33 (d, J = 6.1 Hz, 3H), 0.99 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, MeOD) δ 167.76, 151.82, 136.52, 125.75, 122.40, 117.79, 106.38, 78.82, 30.03, 19.21, 9.67.

3-(*tert*-butoxy)-4-nitrobenzoic acid (2-32c)

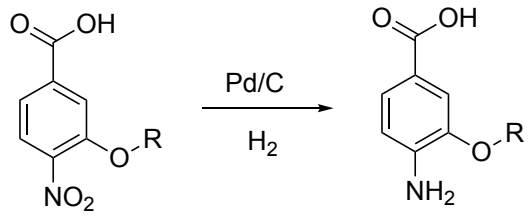


2-31 **2-32c**

3-fluoro-4-nitrobenzoic acid **2-31** (0.5 mmol) was dissolved in 5 ml anhydrous THF in a flame-dried flask under argon and cooled to 0 °C. A 1.0 M solution of potassium tert-butoxide (1.5 mmol) was added dropwise. This solution was stirred and allowed to warm from 0 °C to room temperature overnight and then diluted with ethyl acetate and washed with saturated aqueous NH4Cl. The aqueous layer was back-extracted with CH2Cl2. The organic layers were combined, dried over MgSO4, and concentrated. Purification of the crude mixture by flash chromatography (silica gel, 15:1 CH2Cl2: MeOH) afforded **2-32c** in 90% isolated yield. ^{1}H NMR (400 MHz, MeOD) δ

7.90 (s, 1H), 7.80 (d, J = 8.3 Hz, 1H), 7.74 (d, J = 8.3 Hz, 1H), 1.43 (s, 9H). ^{13}C NMR (101 MHz, MeOD) δ 167.63, 149.76, 125.76, 125.37, 124.96, 84.11, 29.08.

2.4.7. General Procedure for synthesis of 4-aminobenzoic acids (33a-d)



2-32 **2-33**

a: R = Et
b: R = *i*-Pro
c: R = *t*-Bu
d: R = *sec*-Bu

Using 10% Pd/C as catalyst, 50 mL of 4-nitrobenzoic acid (2.0 mmol) in methanol was pumped through the H-Cube®. The pressure of the system was set to full hydrogen mode, and the temperature to 35 °C. The flow rate was 0.8 mL/min. The fraction was analyzed using TLC, which showed complete conversion to the product, and the solvent was reduced to dryness, affording the 4-aminobenzoic acids **2-33a-d**.

4-amino-3-ethoxybenzoic acid (2-33a)

Compound **2-33a** was prepared by the above procedure from 3-ethoxy-4-nitrobenzoic acid **2-32a** in 98% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.57 – 7.30 (m, 2H), 6.71 (d, J = 8.2 Hz, 1H), 4.09 (q, J = 7.0 Hz, 2H), 1.43 (t, J = 7.0 Hz, 3H). ^{13}C NMR (101 MHz, MeOD) δ 170.56, 146.83, 143.50, 125.45, 120.08, 114.36, 113.49, 65.07, 15.11.

4-amino-3-isopropoxybenzoic acid (2-33b)

Compound **2-33b** was prepared by the above procedure from 3-isopropoxy-4-nitrobenzoic acid **2-32b** in 98% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.54 – 7.40 (m,

2H), 6.70 (s, 1H), 4.58 (m, 1H), 1.34 (d, J = 6.0 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 171.77, 145.15, 126.99, 115.39, 113.88, 112.36, 72.35, 26.60, 22.32.

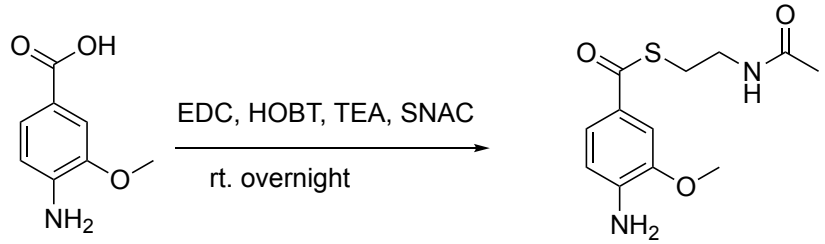
4-amino-3-(*tert*-butoxy) benzoic acid (2-33c)

Compound **2-33c** was prepared by the above procedure from 3-(*tert*-butoxy)-4-nitrobenzoic acid **2-32c** in 99% yield. 1H NMR (400 MHz, MeOD) δ 7.56 (s, 1H), 7.54 (d, J = 8.3 Hz, 1H), 6.71 (d, J = 8.3 Hz, 1H), 1.39 (s, 9H). ^{13}C NMR (101 MHz, MeOD) δ 170.48, 148.85, 142.48, 127.64, 125.04, 119.07, 114.84, 81.01, 29.10.

4-amino-3-(*sec*-butoxy) benzoic acid (2-33d)

Compound **2-33d** was prepared by the above procedure from 3-(*sec*-butoxy)-4-nitrobenzoic acid **2-32d** in 97% yield. ^{1}H NMR (400 MHz, MeOD) δ 7.51 – 7.35 (m, 2H), 6.69 (d, J = 8.2 Hz, 1H), 4.36 (m, 1H), 1.86 – 1.57 (m, 2H), 1.30 (d, J = 6.1 Hz, 3H), 1.00 (t, J = 7.5 Hz, 3H).

S-(2-acetamidoethyl) 4-amino-3-methoxybenzothioate (2-20)

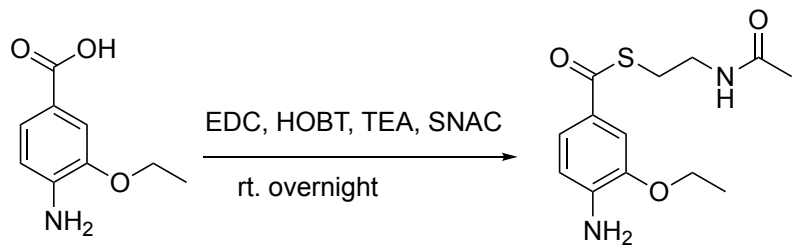


2-34 **2-20**

To a solution of triethylamine (TEA) (2.80 mmol) in dichloromethane (10 mL) was added 4-amino-3-methoxybenzoic acid **2-34** (1.40 mmol), (3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC) (1.40 mmol), 1-hydroxybenzotriazole (HOBt) (1.40 mmol) and N-acetylcysteamine (SNAC) (1.4 mmol). The reaction mixture was stirred overnight under argon. The organic layer was

washed with saturated NaHCO3 solution, 0.1 N HCl solution and brine. It was then dried over anhydrous sodium sulfate and concentrated under vacuum. Purification of the crude mixture by flash chromatography (silica gel, 1:1 to 2:1 of hexanes:EtOAc) afforded **2-20** in 20% isolated yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.53 (d, J = 8.2, 1H), 7.36 (s, 1H), 6.63 (d, J = 8.2 Hz, 1H), 3.89 (s, 3H), 3.51 (dd, J = 12.2, 5.9 Hz, 2H), 3.18 (t, J = 6.3 Hz, 2H), 1.95 (s, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 190.29, 170.29, 146.21, 142.34, 126.60, 122.85, 112.78, 108.72, 55.63, 40.08, 28.31, 23.21.

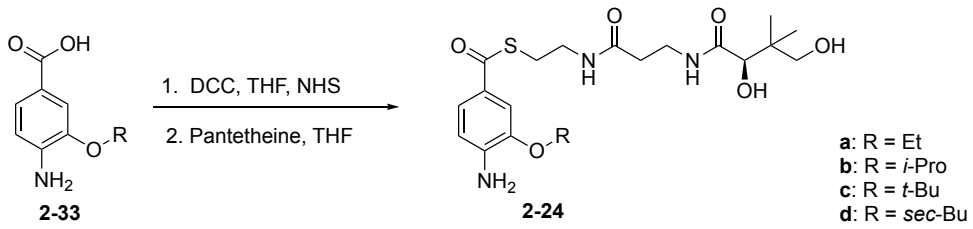
S-(2-acetamidoethyl) 4-amino-3-ethoxybenzothioate (2-21)



2-33a **2-21**

Compound **2-21** was prepared from 4-amino-3-methoxybenzoic acid **2-33a** in the same manner as the above procedure in 30% yield. ^{1}H NMR (400 MHz, CDCl3) δ 7.49 (d, J = 8.2 Hz, 1H), 7.33 (s, 1H), 6.62 (d, J = 8.2 Hz, 1H), 4.09 (q, J = 7.0 Hz, 2H), 3.48 (dd, J = 12.3, 5.9 Hz, 2H), 3.15 (t, J = 6.3 Hz, 2H), 1.93 (s, 3H), 1.42 (t, J = 7.0 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 190.30, 170.29, 145.49, 142.45, 126.56, 122.69, 112.80, 109.63, 77.34, 77.03, 76.71, 64.05, 40.09, 28.30, 23.21, 14.77.

2.4.8. General Procedure for synthesis of S-pantetheinyl-4-aminobenzothioate (2-24a-d)



To a stirred solution of 4-aminobenzoic acid **2-33** (2.0 mmol) in 10 mL of dry THF was added N-hydroxysuccinimide (NHS) (2.0 mmol), followed by N, N'-Dicyclohexylcarbodiimide (DCC) (2.0 mmol) at room temperature. The resulting mixture was stirred overnight at room temperature. The white precipitate was filtered and then washed with EtOAc. The filtrate was concentrated under reduced pressure. The residue was triturated with CH_2Cl_2 to give the NHS ether of **2-33** as a pale brown powder. Pantethine (0.075 mmol), sodium bicarbonate (1.9 mmol), and 1,4-dithiothreitol (DTT) (0.08 mmol) were dissolved in 3 mL of water, and left to stand for 10 minutes to allow reduction of the pantethine to pantetheine to take place. The NHS ester (0.54 mmol) was dissolved in 7 mL of tetrahydrofuran (THF) and added to the 3 mL water mixture. The reaction mixture was stirred overnight and then the solvent was evaporated under reduced pressure. The crude product was then dissolved in water and the insoluble residue was removed by filtration. The resulting solution was lyophilized and purified by flash chromatography (silica gel, 20:1 to 10:1 CH_2Cl_2: MeOH). The fraction was further purified by Prep HPLC using a SPLC-18DB column (10 x 250 mm, 5 µm, Supelco) at a flow rate of 2 mL/min on an Agilent 100 HPLC system with a quaternary pump and an automatic injector.

(R)-S-pantetheinyl-4-amino-3-methoxybenzothioate (2-33)

Compound **2-33** was prepared by the above procedure from 4-amino-3-methoxybenzoic acid **2-34**. ^{1}H NMR (400 MHz, D_2O) δ 7.52 (dd, J = 8.3, 1.9 Hz, 1H), 7.35 (d, J = 1.9 Hz, 1H), 6.79 (d, J = 8.3 Hz, 1H), 3.91 (s, 1H), 3.88 (s, 3H), 3.47 – 3.29 (m, 6H), 3.17 (t, J = 6.3 Hz, 2H), 2.44 (t, J = 6.5 Hz, 2H), 0.83 (d, J = 13.9 Hz, 6H). ^{13}C NMR (101 MHz, D_2O) δ 193.17, 175.04, 174.02, 146.63, 126.36, 123.04, 114.04, 109.43, 75.86, 68.49, 55.92, 38.93, 38.58, 35.50, 35.33, 28.03, 22.30, 20.46, 19.11.

(R)-S-pantetheinyl-4-amino-3-ethoxybenzothioate (2-24a)

Compound **2-24a** was prepared by the above procedure from 4-amino-3-ethoxybenzoic acid **2-33a**. ^{1}H NMR (400 MHz, D_2O) δ 7.56 (dd, J = 8.3, 1.9 Hz, 1H), 7.41 (d, J = 1.8 Hz, 1H), 6.84 (d, J = 8.3 Hz, 1H), 4.15 (q, J = 7.0 Hz, 2H), 3.91 (s, 1H), 3.54 – 3.25 (m, 6H), 3.20 (t, J = 6.2 Hz, 2H), 2.44 (t, J = 6.5 Hz, 2H), 1.40 (t, J = 7.0 Hz, 3H), 0.83 (d, J = 13.8 Hz, 6H). 13C NMR (101 MHz, D_2O) δ 193.40, 175.05, 174.08, 145.68, 143.76, 126.41, 123.11, 114.25, 111.11, 75.86, 68.48, 65.24, 38.90, 38.58, 35.51, 35.34, 28.06, 20.45, 19.09, 14.02.

(R)-S-pantetheinyl-4-amino-3-isopropoxybenzothioate (2-24b)

Compound **2-24b** was prepared by the above procedure from 4-amino-3-isopropoxybenzoic acid **2-33b**. ^{1}H NMR (400 MHz, D_2O) δ 7.54 (dd, J = 8.4, 1.8 Hz, 1H), 7.45 (d, J = 1.8 Hz, 1H), 6.82 (d, J = 8.4 Hz, 1H), 4.62 (m, 1H), 3.88 (s, 1H), 3.48 – 3.21 (m, 6H), 3.17 (t, J = 6.2 Hz, 2H), 2.41 (t, J = 6.5 Hz, 2H), 1.30 (d, J = 6.1 Hz, 6H), 0.79 (d, J = 14.0 Hz, 6H). ^{13}C NMR (101 MHz, D_2O) δ 193.36, 175.05, 174.07, 144.96,

144.19, 126.43, 123.38, 114.65, 114.13, 75.86, 73.18, 68.48, 38.90, 38.59, 35.50, 35.35, 28.07, 21.21, 20.47, 19.11

(R)-S-pantetheinyl-4-amino-3-(*tert*-butoxy)benzothioate (2-24c)

Compound **2-24c** was prepared by the above procedure from 4-amino-3-(*tert*-butoxy)benzoic acid **2-33c**. ^{1}H NMR (400 MHz, MeOD) δ 7.65 (dd, J = 8.4, 2.0 Hz, 1H), 7.59 (d, J = 2.0 Hz, 1H), 6.88 (d, J = 8.5 Hz, 1H), 3.92 (s, 1H), 3.49 – 3.32 (m, 6H), 3.20 (t, J = 6.2 Hz, 2H), 2.44 (t, J = 6.5 Hz, 2H), 1.41 (s, 9H), 0.83 (d, J = 14.1 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 193.19, 175.05, 174.06, 148.38, 140.87, 125.93, 125.41, 122.52, 115.25, 82.68, 75.84, 68.48, 38.89, 38.61, 35.51, 35.36, 28.08, 27.86, 20.50, 19.14.

(R)-S-pantetheinyl-4-amino-3-(*sec*-butoxy)benzothioate (2-34d)

Compound **2-22d** was prepared by the above procedure from 4-amino-3-(*sec*-butoxy)benzoic acid **2-33d**. ^{1}H NMR (400 MHz, MeOD) δ 7.48 (dd, J = 8.3, 1.8 Hz, 1H), 7.38 (d, J = 1.7 Hz, 1H), 6.71 (d, J = 8.3 Hz, 1H), 4.40 (m, 1H), 3.92 (s, 1H), 3.58 – 3.37 (m, 6H), 3.16 (t, J = 6.6 Hz, 2H), 2.44 (t, J = 6.6 Hz, 2H), 1.82-1.64 (m, 2H), 1.33 (d, J = 6.1 Hz, 3H), 1.03 (t, J = 7.5 Hz, 3H), 0.94 (s, 6H). ^{13}C NMR (101 MHz, MeOD) δ 189.63, 174.62, 172.51, 144.87, 144.07, 125.29, 122.32, 112.41, 111.11, 75.98, 75.79, 69.01, 39.11, 38.95, 35.03, 34.36.95, 28.76, 27.51, 19.89, 19.53, 18.15, 8.63.

2.4.9. Synthesis of AMB-CoA analog 2-27

4-amino-3-methoxybenzoic acid **2-34** (1.4 mg, 8.7 μmol), coenzyme A (sodium salt hydrate, 9 mg, 13 μmol), and benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP, 6.7 mg, 13 μmol) were dissolved in 500 μL of a

tetrahydrofuran/ddH2O (1:1, v/v). N,N-Diisopropylethylamine (DIPEA, 15 µL, 87 µmol) was added, the reaction mixture was stirred at 25 °C for 2 h. and then lyophilized. The dried product was resuspended in 10% aqueous acetonitrile and purified by preparative HPLC. The product-containing fractions were combined and lyophilized yielding AMB-CoA **2-27** as a white solid. ^{1}H NMR (400 MHz, D_2O) δ 8.45 (s, 1H), 8.10 (s, 1H), 7.37 (dd, J = 8.3, 1.6 Hz, 1H), 7.21 (d, J = 1.6 Hz, 1H), 6.70 (d, J = 8.3 Hz, 1H), 6.06 (d, J = 5.6 Hz, 1H), 4.75, 4.55 (m, 1H), 4.23 (m, 2H), 4.00 (s, 1H), 3.85 (s, 3H), 3.82 – 3.43 (m, 6H), 3.18 (m, 2H), 2.43 (m, 2H), 0.89 (s, 3H), 0.75 (s, 3H). ^{13}C NMR (101 MHz, D2O) δ 192.92, 174.78, 174.09, 155.40, 152.65, 146.42, 143.10, 139.66, 126.07, 122.86, 113.91, 108.95, 86.60, 74.01, 71.97, 55.84, 38.96, 38.43, 35.44, 28.04, 23.28, 20.94, 18.25.

2.4.10. Overexpression and purification of CysG1

CysG1 (adenylation domain and peptidyl-carrier-protein of CysG module 5, (Baumann et al., 2014) was cloned into a pHisTEV plasmid (derived from pET30a). Single colony transformants of *E. coli* BL21(DE3) carrying pHisTEV-CysG1 were grown aerobically for 12-15 h at 37 °C in 10 mL of LB medium supplemented with 50 µg/mL kanamycin (LB-kan). Cells from the starter culture were added to 1 L of LB-kan, and the culture was grown aerobically at 37 °C until the A600 was 0.6-0.8. Expression was induced by the addition of 1 mM IPTG and the culture was grown aerobically at 15 °C for an additional ~ 20 h. Cells were harvested by centrifugation. The medium was discarded and the cell paste was stored at –80 °C until use. Cells were thawed, resuspended in 20 mL of Lysis Buffer (50 mM potassium phosphate, pH 7.8, 150 mM

NaCl), and then lysed by sonication on ice for 6 × 20 s using a Misonix Sonicator 3000 (24 W, 1.5 s pulse, 1.5 s pause). Cell debris was removed by centrifugation and the lysate was loaded onto a Histrap column pre-equilibrated with lysis buffer. The column was washed with 5-10 column volumes of wash buffer (100 mM Tris-HCl, 30 mM Imidazole, pH 7.5). The protein was then eluted using elution buffer (100 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH 7.5). Brown colored fractions containing the desired protein were pooled and buffer exchanged to final buffer (100 mM Tris-HCl, pH 7.5 at 25 °C, 30% glycerol) via Econo-Pac 10DG size exclusion chromatography. The purified enzyme was stored in liquid nitrogen. Protein concentration was measured by the absorbance at 280 nm (A280) with an extinction coefficient calculated using the ProtParam tool of the ExPASy proteomics server.

2.4.11. Chemoenzymatic synthesis of the 4-amino-3-methoxybenzoate (AMB) CysG1 thioester 2-28

A typical enzymatic reaction for synthesizing the AMB CysG1 thioester **2-28** was performed in 100 mM phosphate buffer, pH 7.5 containing the Svp phosphopantetheinyl transferase (20 μM), CysG1 (43 μM), AMB-CoA (0.3 mM) and $MgCl_2$ (10 mM). The reaction was allowed to proceed aerobically at room temperature for about 6 hrs and was quenched by desalting using a Bio-spin 6 column (Bio-Rad). Then the sample was injected onto a reverse phase HPLC Zorbax 300SB-C8 3.5μ, 3×150 mm column operating at room temperature coupled to an ESI-TOF mass spectrometer operating in positive ionization mode. LC-MS analysis of the reaction mixture demonstrated the expected around 490 Daltons mass increase to CysG1

resulting from the formation of the phosphopantetheinyl-AMB thioester **2-28** (Figure 2.18).

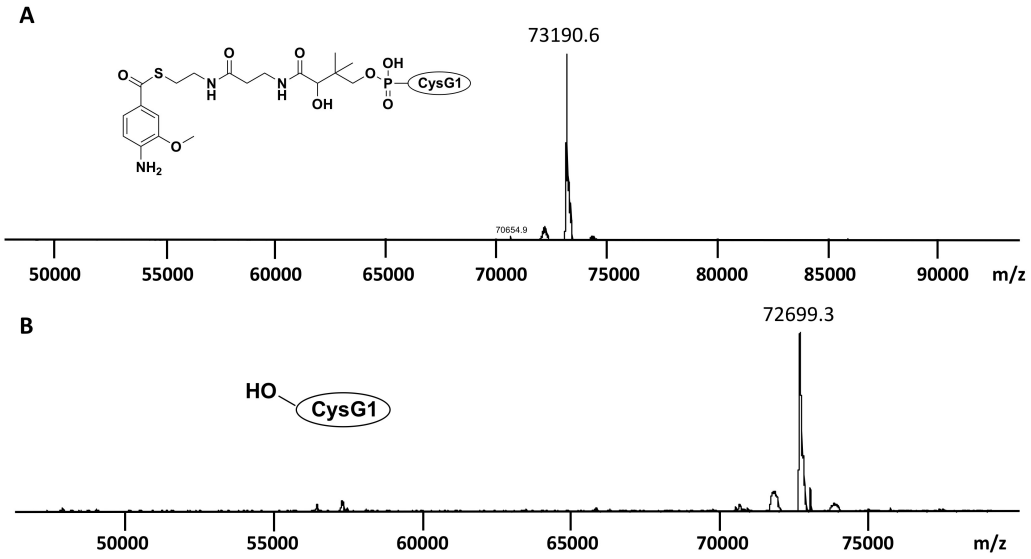


Figure 2.18 MS analysis of AMB-CysG1 thioester **2-28**. A: CysG1 thioester **2-28** (calculated monoisotopic mass: 73189.7 Da, experimental mass: 73190.6 Da) and B: CysG1 without modification (calculated monoisotopic mass: 72700.5 Da, experimental mass: 72699.3 Da).

2.4.12. Reconstitution of the CysS-catalyzed methyl transfer reaction

AMB-CysG1 thioester (53 μM), prepared as described above, was buffer exchanged into anaerobic 100 mM phosphate buffer, pH 7.5 using a Bio-spin 6 column in a glove box and then treated with 0.28 mM MeCbl, 0.56 mM SAM and 0.6 mM methyl viologen, 1.8 mM NADPH and 22 μM CysS or 44 μM CysS (The formation of trimethylated products required higher enzyme concentrations). The reaction was

allowed to proceed anaerobically at room temperature for about 14 hrs. The protein in the reaction mixture was desalted using 30K Amicon Ultra-0.5 mL centrifugal filters.

LC-ESI-TOF-MS on protein samples was performed using an Agilent 1260 HPLC system equipped with a binary pump and a 1200 series diode array detector followed by a MicroToF-Q II mass spectrometer (Bruker Daltonics) using an ESI source in positive mode. The analysis was performed on a Zorbax 300SB-C8 column (3 ×150 mm column, 3.5 μm particles, Agilent).

LC conditions: A- 0.1 % formic acid in H_2O, B- 0.1 % formic acid in acetonitrile

LC method: 0 min – 90% A, 10 min – 80% A, 25 min – 10% A 90% B, 27 min – 90% A, 30 min – 90% A.

The mass spectrometer parameters for protein analysis: Capillary, 4500 V; end plate offset, 500 V; nebulizer gas, 3.0 bar; dry gas, 10.0 L/min; dry gas temperature, 220°C; funnel 1 RF, 400.0 Vpp; funnel 2 RF, 250 Vpp; quadrupole, Ion energy, 4.0 eV; low mass, 200 m/z; collision cell, collision energy, 8.0 eV; collision RF, 1000.0 Vpp; transfer time, 120.0 μs; prepulse storage, 10.0 μs. Data were processed with DataAnalysis ver. 4.1 (Bruker Daltonics)

3. MECHANISTIC STUDIES OF SEQUENTIAL METHYLTIONS BY CysS*

3.1. Introduction

Methylation has played a crucial role in many biological processes. The classic nucleophilic substitution is well established to generate the methylated product. There is not much known for the methylations at non-nucleophilic centers in numerous natural products, which have been recently found to be catalyzed by the enzymes in the radical SAM superfamily. Among them, cobalamin-dependent radical SAM methyltransferases can catalyze unusual methylations on diverse classes of substrates, including sp^2-hybridized, sp^3-hybridized carbon atoms, and phosphinate phosphorous atoms.[20] Although, they are widespread in nature and have been first discovered in 1995,[58] it was not until recently more and more enzymes in this class have been reconstituted *in vitro*, such as PhpK for the phosphinic acid P-methylation[59], GenK and Fom3 for the alcohol C-methylation[26, 28], ThnK for the iterative C-methylation to form the ethyl group[30], TsrM for the aromatic C-methylation[32], PoyC for the valine residue C-methylation[33] (Figure 3.1). However, details of mechanistic studies are still hindered because of the poor activity of the protein and substrates complexity.

---

* Wang, Y.; Begley, T. P. Mechanistic studies of cobalamin dependent radical SAM enzyme CysS. *Manuscripts in Preparation.*

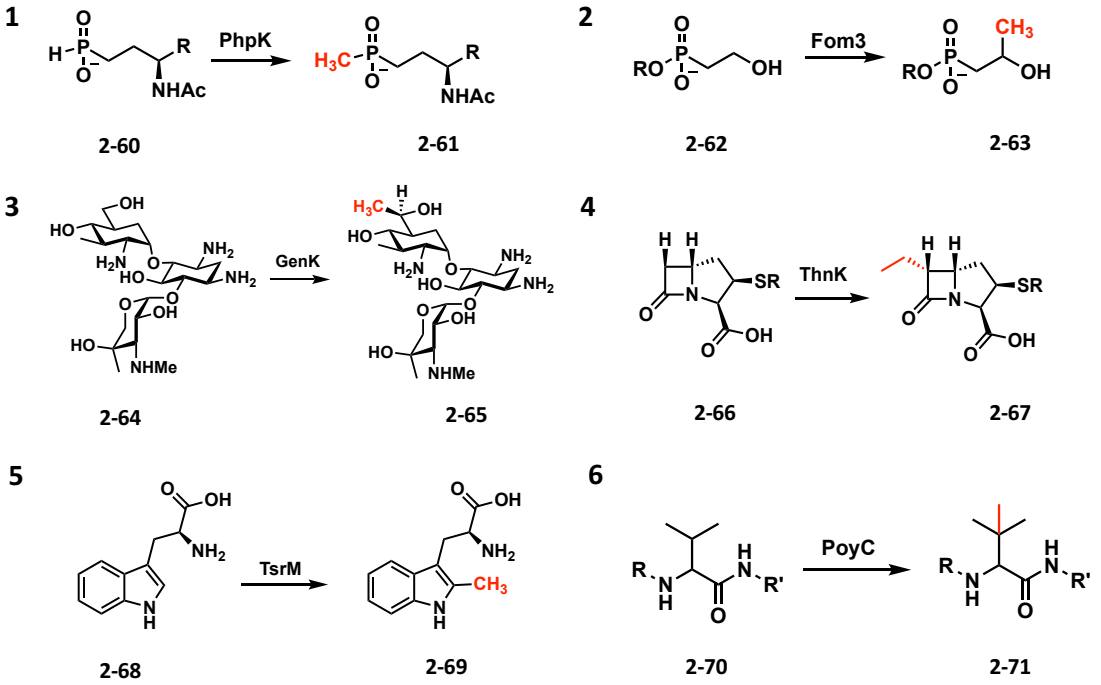


Figure 3.1 Selected secondary metabolites whose biosynthesis include class B RSMTs that have been reconstituted *in vitro*. Carbon atoms that are introduced by class B RSMTs are highlighted by red.

After it is shown that CysS can perform sequential methylations on a methyl group to form ethyl, isopropyl, *sec*-butyl, and *t*- butyl groups on *p*-aminobenzoate thioester by a radical mechanism *in vitro*, other combinations on methyl group of *p*-aminobenzoate thioester in cystobactamids analogs were also explored. At first, only isopropoxylation was discovered on the 3-hydroxy-4-aminobenzoic acid moieties in the antibiotic cystobactamids from *Cystobacter sp*.[51] Later, ethyl and *sec*-butyl groups in addition to the isopropyl group were also found in the new cystobactamids derivatives. [60] Some of these new derivatives display significantly improved antibacterial activities in the low μM range against Gram-negative pathogens.[60] Studying the mechanism of how

these modifications occur in nature will help us better understand the structure–activity relationships of antibiotic discovery.

A mechanism is proposed for CysS in Figure 3.2. After initial formation of methylcobalamin by SAM mediated methylation, a second SAM is reductively cleaved by the $[4Fe-4S]^{+1}$ cluster to generate the 5′-dA •. This 5′-dA • abstracts a hydrogen atom from the methyl group of the substrate **2-23** to give the substrate radical **2-25**, which then undergoes a radical substitution with methylcobalamin to give the ethyl ether **2-24a**. Regeneration of MeCbl from Cbl(II) can be achieved by reduction to Cbl(I) by the $[4Fe-4S]^{+1}$ cluster followed by SAM-mediated methylation. Repetition of this sequence results in the successive formation of the isopropyl, *t*-butyl, and *sec*-butyl ethers of **2-23**.

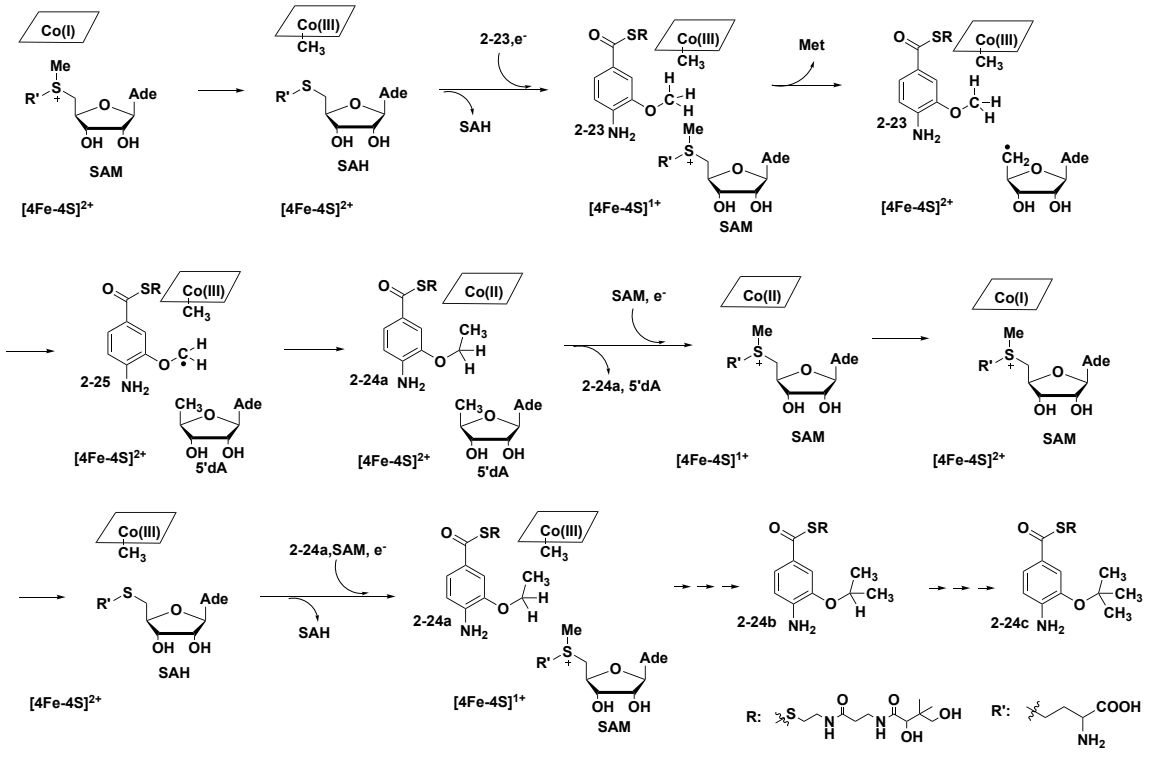


Figure 3.2 Proposed mechanism for CysS-catalyzed iterative methylations.

Because of the poor activities from the inclusion bodies and substrate complexities, detailed mechanism studies were not available for B_{12}-dependent radical SAM methyltransferase, except for few stereochemistry studies[26-29]. Here, we are able to use a radical clock probe to trap the substrate radical and estimate the radical substitution of a methyl group from cobalamin in the mechanism. A brominated substrate analog is designed to trap the Cbl(I) intermediate, which is the critical species for the regeneration of MeCbl during each turnover. The resulting crosslinking species also provides useful information on structural orientations of cofactors in the enzyme active site. Competitive

and intrinsic isotope effects were also obtained to investigate the transition step of hydrogen abstraction step during the reaction.

3.2. Result and discussion

3.2.1. Measure methyl transfer rate with a radical clock probe

During the CysS catalytic cycle, it is proposed that the substrate radical can undergo a radical substitution with methylcobalamin to give the methylated product. Since all the classic B_{12}-dependent methyltransferases go through nucleophilic attack of the substrate to get methylated, the racial mechanism proposal was considered speculative for a long time. Even though such methyl group transfer from MeCbl to a carbon-based radical exists in cobalamin model studies and radical substitutions are widely used in organic synthesis, there is no such precedent of a radical species get methylated by MeCbl for the B_{12}-dependent methyltransferases in nature until more and more B_{12}-dependent radical SAM enzymes were discovered recently. For GenK, both nucleophilic and radical mechanism are proposed for the methyl group transfer from the MeCbl to the substrate and currently there is no experimental evidence that can distinguish them. But for ThnK, Fom3, and PoyC, it's more likely that radical mechanism is responsible for the methylation of the substrates.

In order to probe the radical substitution step for methylation, a cyclopropyl analog **2-41** was designed. The ring opening by the cyclopropylcarbinyl radical is a well-known fast radical reaction, which has been used as a radical trap and a measurement of the unknown rate in many competition reactions. The reported cyclopropylcarbinyl radical ring opening rate constant of (methoxymethyl) –cyclopropane **2-47** is $2 \times 10^7\ s^{-1}$.

When cyclopropyl analog was incubated with CysS, SAM, MeCbl, MV and NADPH, LC-MS analysis of the CysS reaction mixture showed the formation of one of the new products had a mass corresponding to the methylated product **2-43**, which was confirmed by the synthetic standard (Figure 3.3A, B). A large amount of 5′-dA formation was also observed in the same reaction when compared with the no substrate control, which further demonstrated the cyclopropyl analog was accepted as a substrate for CysS. When the reaction mixture was treated with 1% TFA, the ring opening derived product aminophenol compound **2-46** was detected by LC-MS (Figure 3.3C). The ratio of the methylated compound **2-43** and aminophenol compound **2-46** was quantified by using the authentic standard as 12.0. If we use the literature reported value (2×10^7 s^{-1}) to estimate the value of the ring opening rate constant of the cyclopropyl substrate **2-41** (k_{ring}), we can obtain the rate constant of the methyl group transfer from MeCbl to substrate radical (k_{me}) based on the ratio of the final products **2-43** and **2-46**. The estimated k_{me} is around 2.4×10^8 s^{-1}, which has never been estimated before. Since the cyclopropyl analog is a methoxybenzothioate derivative instead of a simple methoxy group as in the reference reaction and there might be a conformational restriction at the enzyme active site to create stereoelectronic effects, the exact ring opening rate constant of cyclopropyl analog would be lower than 2×10^7 s^{-1}. Thus, the rate constant of the methyl group transfer by radical substitution could be lower than 2.4×10^8 s^{-1}. Nevertheless, the methyl transfer rate constant is comparable with the ring opening rate constant of cyclopropyl analog **2-43**. In solution chemistry, it's reported that a ((6-iodo-2-methylhexan-3-yl)oxy)diphenyl(trimethylstannyl)silane species can go through an

intramolecular homolytic substitution reaction at silicon (S_Hi) to form a cyclic five-membered alkoxysilanes and a trimethylstannyl radical at a rate constant as 1×10^7 s^{-1}.[61] The radical substitution rate constant could be larger when the bond orientation is restrained in the enzyme active site.

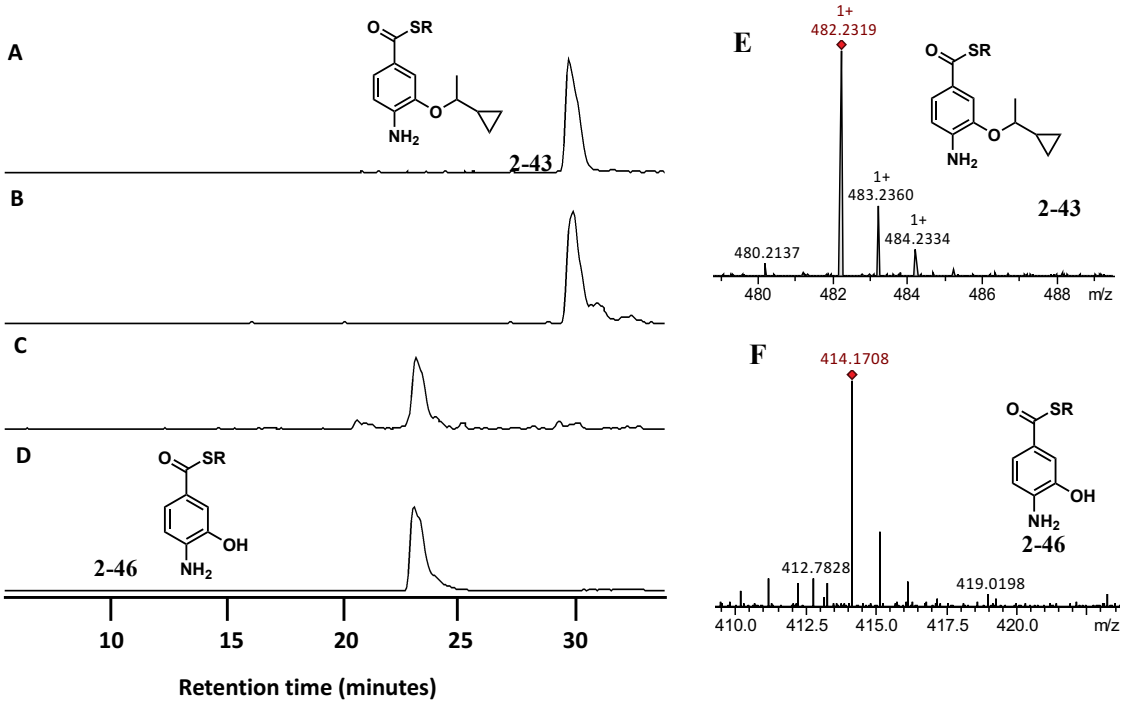


Figure 3.3 CysS catalyzed reaction with cyclopropyl analog, extracted ion chromatograms (EICs) of the methylated cycropropyl product **2-43** $[M + H]^+$ (482.23 ± 0.02) for (A) Synthetic standard; (B) CysS catalyzed reaction; EICs of the ring opening derived product **2-46** $[M + H]^+$ (414.17 ± 0.02) for (C) CysS catalyzed reaction; (D) Synthetic standard; (E) Mass spectrum of the methylated cyclopropyl product **2-43** in CysS catalyzed reaction (Exact mass: 482.2319, Accuracy: <1 ppm); (F) Mass spectrum of the ring opening derived product **2-46** in CysS catalyzed reaction (Exact mass: 414.1693, Accuracy: 3.6 ppm); (G) Strategy for measuring the methyl transfer rate using the ring opening rate of cyclopropylcarbinyl radical.

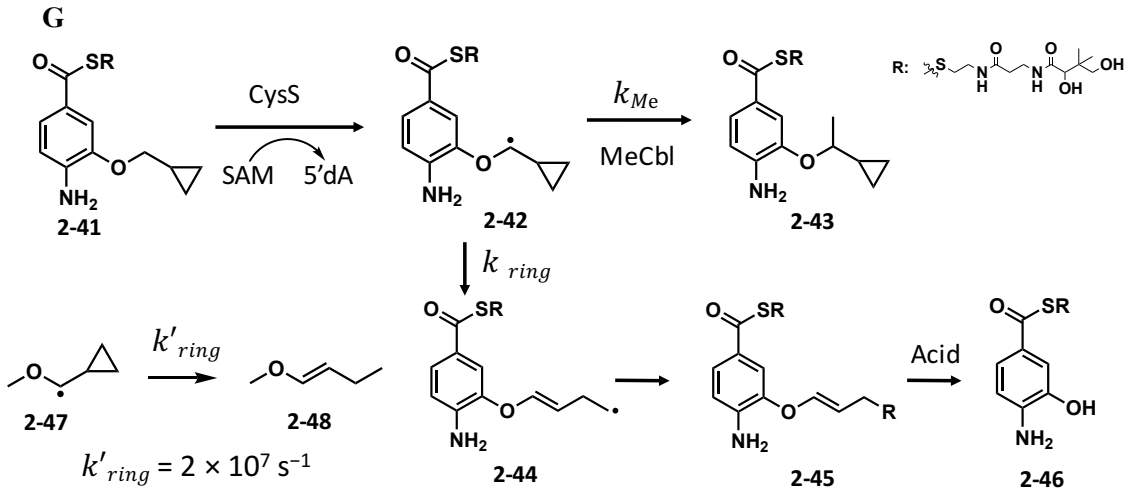


Figure 3.3 continued

3.2.2. Trapping a Cbl(I) intermediate by brominated substrate analog

Since the rate of the methyl transfer from Cbl to the substrate is very fast, we can expect that the carbon atom from the methoxyl group of the substrate should be in close vicinity of the carbon from the MeCbl. As we proposed Cbl(I) species is formed for the regeneration of the MeCbl (Figure 3.1) and Cbl(I) is known as a super nucleophile, if we replace the hydrogen on the methoxy carbon with a good leaving group, like a bromide, Cbl(I) can have a nucleophilic substitution to form a cobalamin substrate cross-linking species (Figure 3.4A). In order to explore this idea, brominated substrate analog **2-49** was synthesized and was first tested in the model reaction to investigate whether Cbl(I) can react with brominated substrate analog **2-49** to generate the cross-linking species. Cbl(I) was first chemically reduced from Cbl(III) by using an excess of $NaBH_4$ anaerobically, indicated by the UV spectrum (Figure 3.5). When brominated substrate analog **2-49** was added to the reduced Cbl(I) reaction mixture, the color immediately

changed from light gray to pink, which suggested a Cbl(III) species was formed. Further HPLC and LCMS analysis of the reaction mixture revealed the new species had a similar UV spectrum of Cbl(III) and had a MS corresponding to the Cbl crosslinking species **2-50** (Figure 3.6).

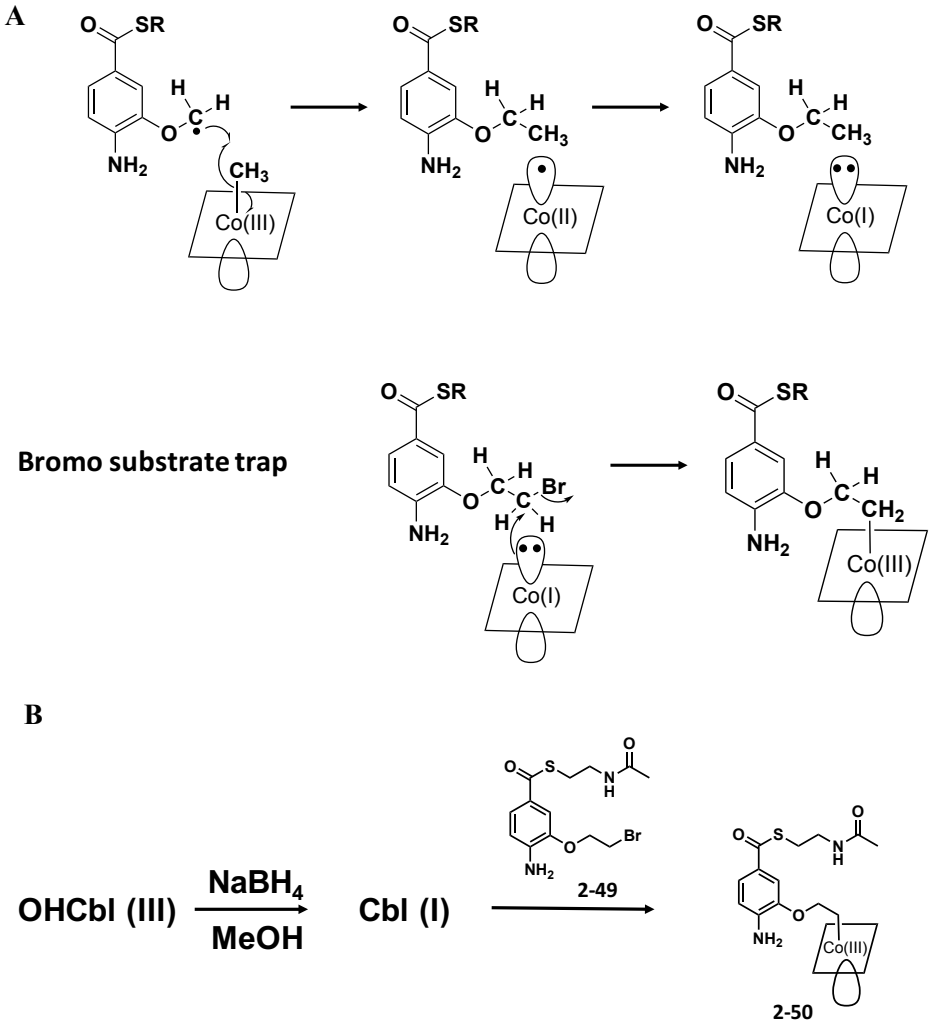


Figure 3.4 (A) Strategy of developing the brominated substrate analog to trap Cbl(I) species (B) Model reaction for cobalamin cross-linking species through Cbl(I)

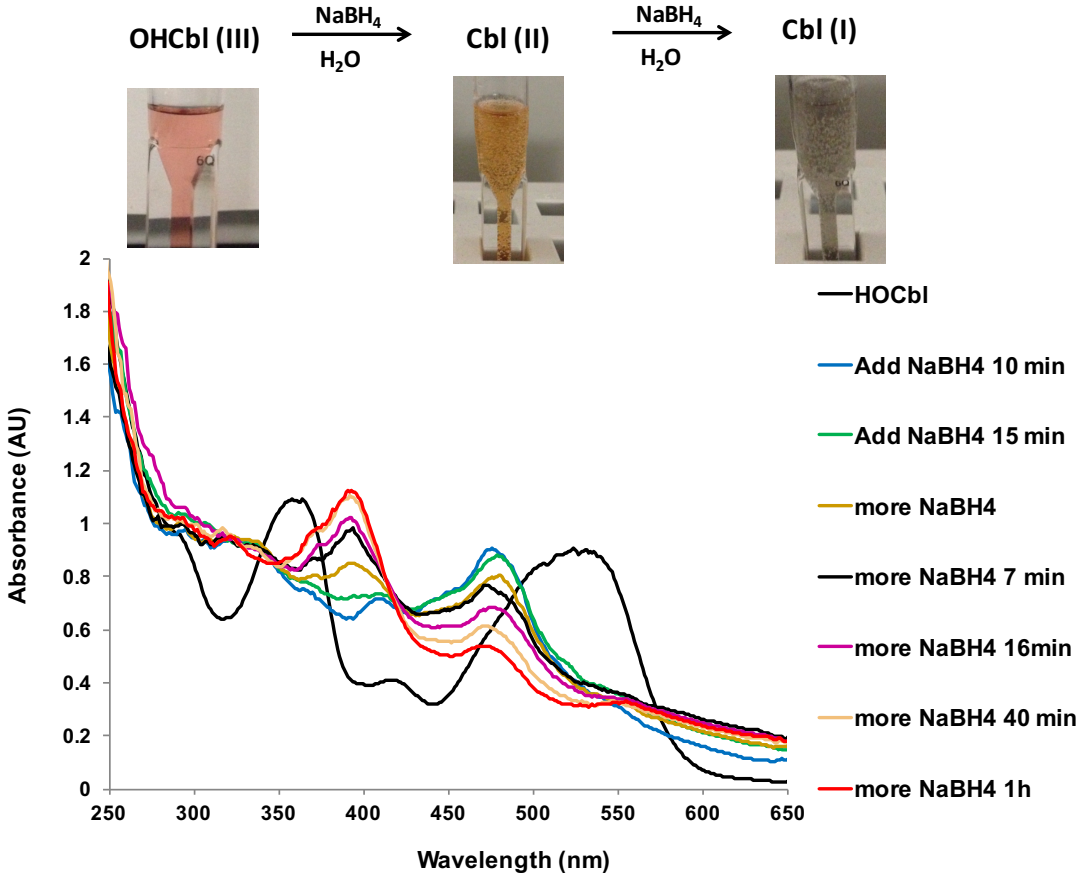


Figure 3.5 UV spectrum of monitoring the chemical reduction of from Cbl(III) to Cbl(I)

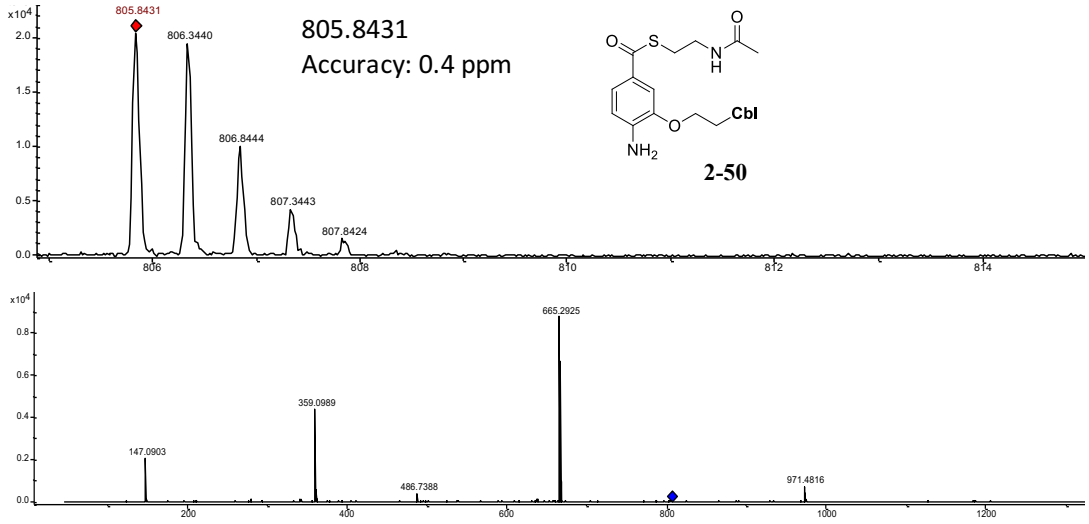


Figure 3.6 Upper panel: MS of the substrate cobalamin adduct **2-50** in the model reaction; lower panel: MS/MS of the substrate cobalamin adduct **2-50** in model reaction

Then, brominated substrate analog **2-49** was incubated with CysS, OHCbl, MV, NADPH, a new peak was detected by LC-MS, which has a MS corresponding with the Cbl substrate crosslinking species **2-50** (Figure 3.7B). This peak is absent when CysS, Cbl, **2-49** was excluded in the enzymatic reaction (Figure 3.7A). Also, the new peak had the same retention time, MS and MS/MS spectrum as the compound **2-50** formed in the model reaction (Figure 3.8, 3.9).

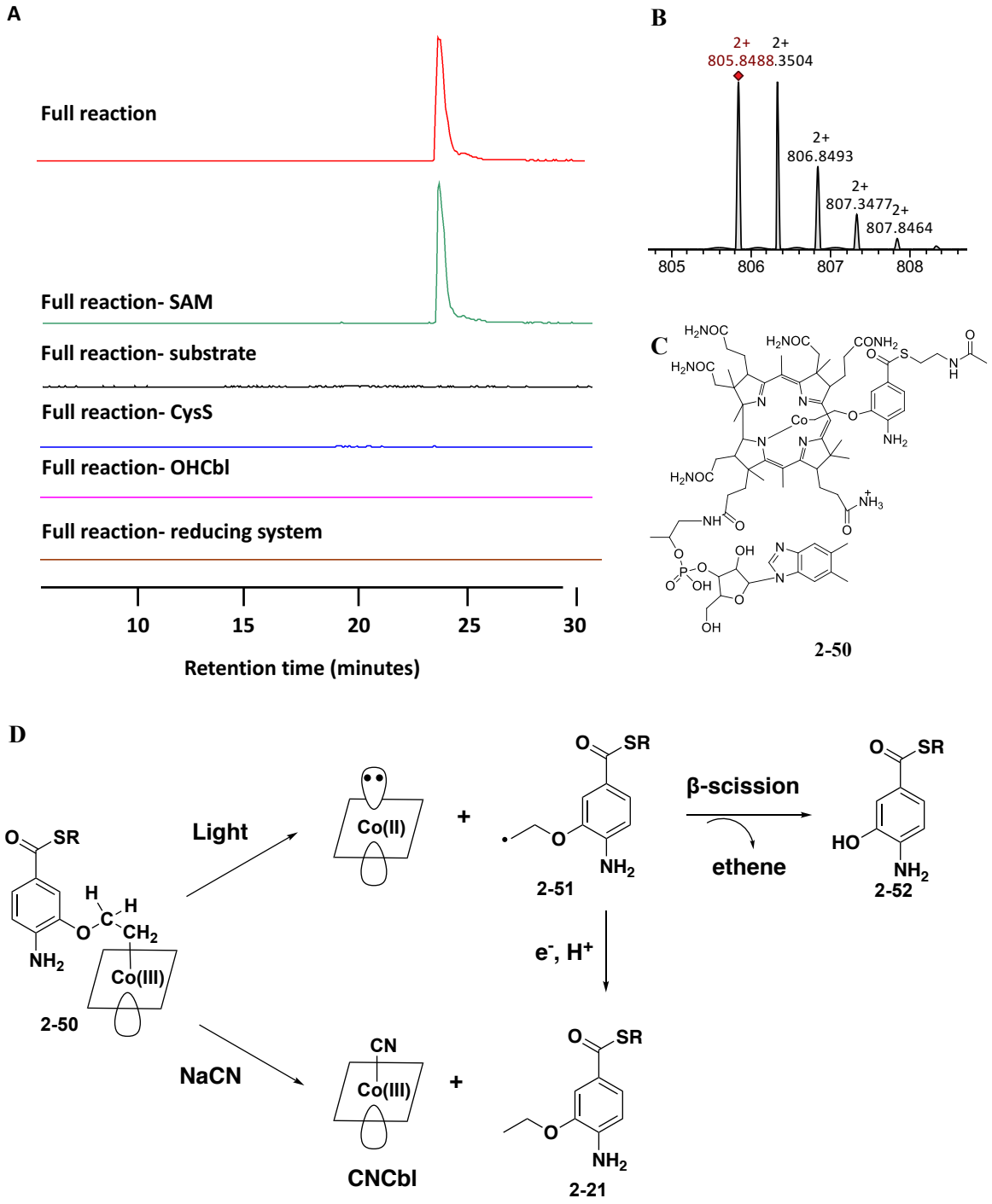


Figure 3.7 CysS-catalyzed reaction with brominated analog **2-49**. (A) LCMS analysis of the CysS reaction with brominated analog. EICs of the adduct **2-50** $[M + 2H]^{2+}$ (805.84 ± 0.02), red trace represents full reaction; green, black, blue, pink and black trace represent the full reaction without SAM, substrate, CysS, OHCbl and reducing system, respectively. (B) Mass spectrum of the adduct **2-50** in CysS catalyzed reaction (Exact mass: 805.8464, Accuracy: 9.4 ppm). (C) Structure of the Cbl substrate crosslinking species 2-50 (D) Different products generated from the crosslinking species under light condition or with NaCN.

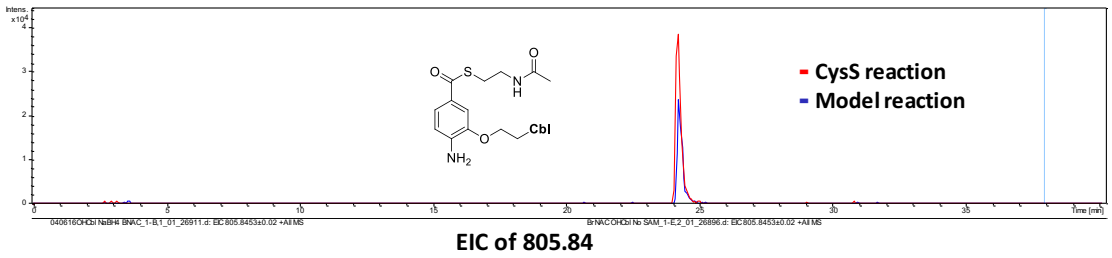


Figure 3.8 LC-MS detection of the CysS-catalyzed reaction with brominated analog **2-49** and the model reaction. EICs of the substrate cobalamin crosslinking adduct **2-50** $[M + 2H]^{2+}$ (805.84 ± 0.02).

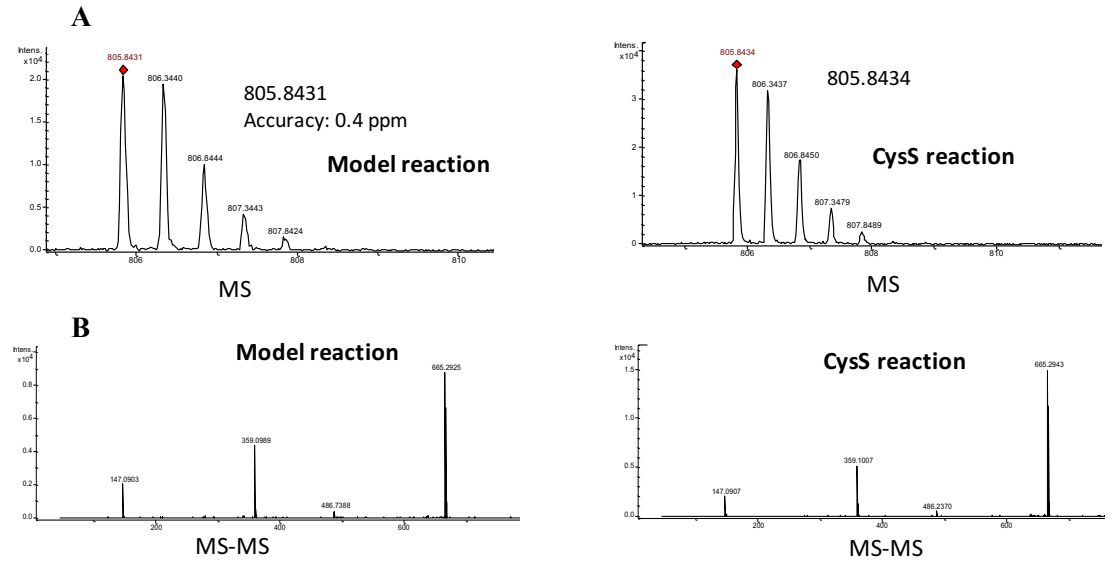


Figure 3.9 LC-MS detection of the CysS-catalyzed reaction with brominated analog **2-49** and the model reaction. (A) MS spectrum at 24 min. (B) Mass fragments of 805.84.

To further confirm the formation of the carbon cobalt covalent bond, the enzymatic reaction was exposed to the light (Figure 3.7D). After light treatment, the

amount of the substrate Cbl crosslinking species **2-50** decreased while the intensity of several peaks was increased (Figure 3.10).

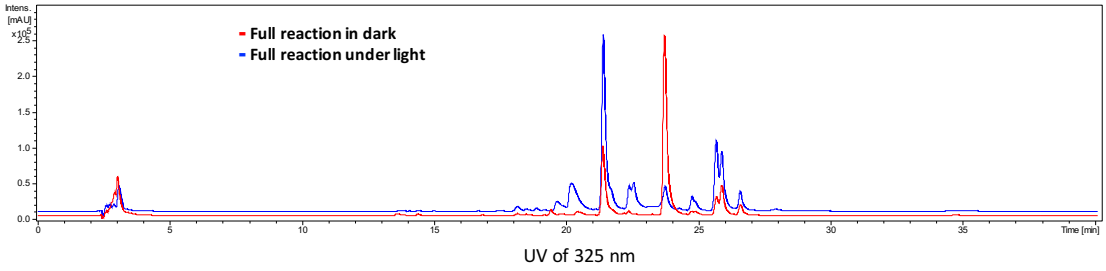


Figure 3.10 UV spectrum of CysS catalyzed reaction with (blue) and without (red) exposure to light.

By co-elution experiment with the synthetic standard, the peak at 21 minute corresponds to aminophenol compound **2-52** and the peak at 25 minute corresponds to the ethyl compound **2-21**(Figure 3.11). Since carbon cobalt is considered weak, under the light condition, the carbon cobalt bond can go through homolytic cleavage to form ethyl radical **2-51**, which either have a beta scission to form aminophenol compound **2-52** or abstract a hydrogen atom to form the ethyl compound **2-21**(Figure 3.7D). In order to test whether the radical abstract a hydrogen atom from the solvent, the reaction was performed in D_2O. Compared with the H_2O buffer, M+1 peak of **2-21** has increased from 19% to 65% of the M peak (Figure 3.12).

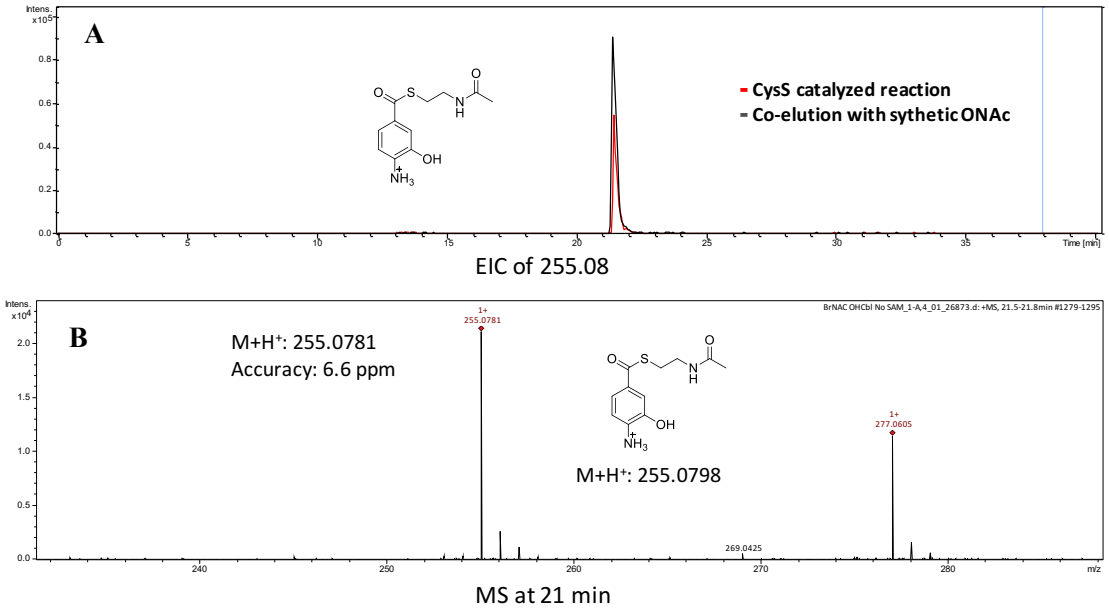


Figure 3.11 LC-MS detection of **2-52** in the CysS reaction using brominated analog **2-49** as substrate. (A) EICs of the aminophenol compound **2-52** $[M + H]^+$ (255.08 ± 0.02). (B) Mass spectrum of the aminophenol compound **2-52**.

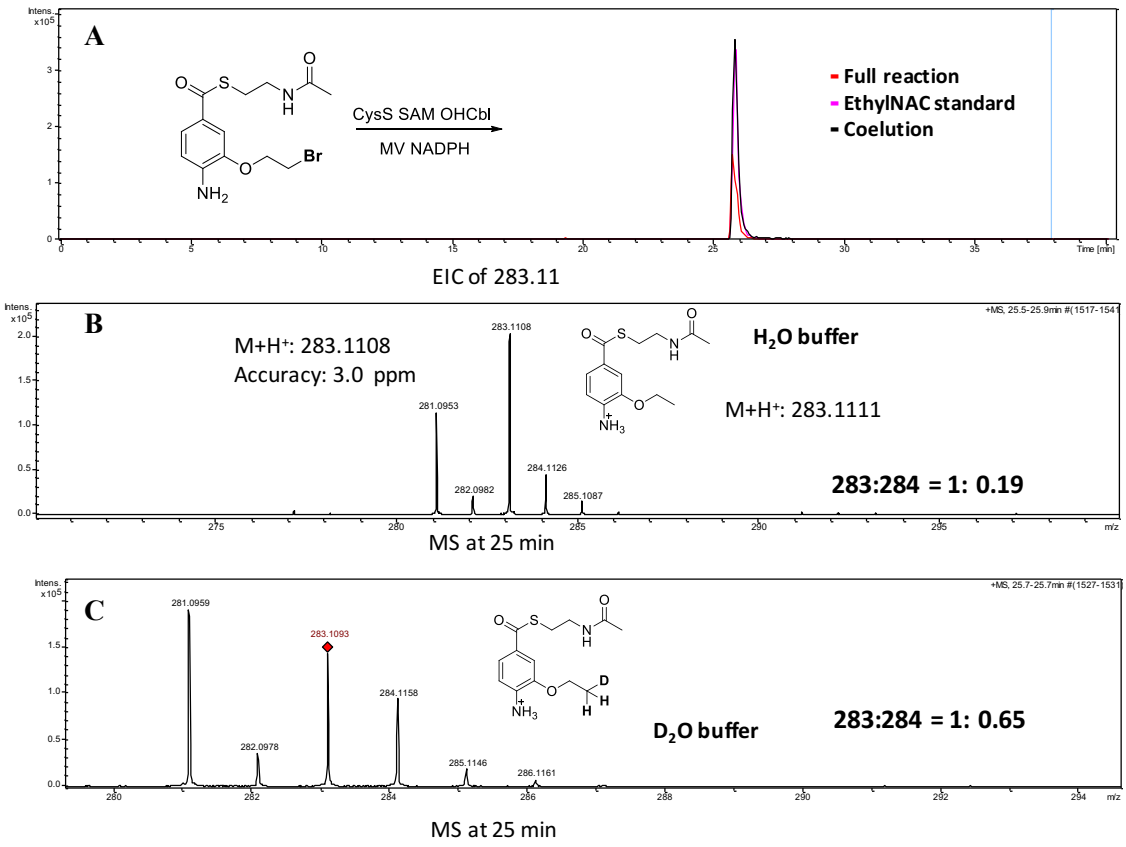


Figure 3.12 LC-MS detection of the ethyl compounds in CysS reaction using brominated analog **2-49** as substrate. (A) EICs of the ethyl compound **2-21** $[M + H]^+$ (283.11 ± 0.02). (B) Mass spectrum of the ethyl compound **2-21** in H_2O buffer. (C) Mass spectrum of the ethyl compound in D_2O buffer

The reaction mixture was also derivatized by adding cyanide (Figure 3.7D). Complete consumption of the crosslinking species **2-50** was detected with the formation of the aminophenol compound **2-52** and cyanide cobalamin (CN-Cbl) by HPLC (Figure 3.13).

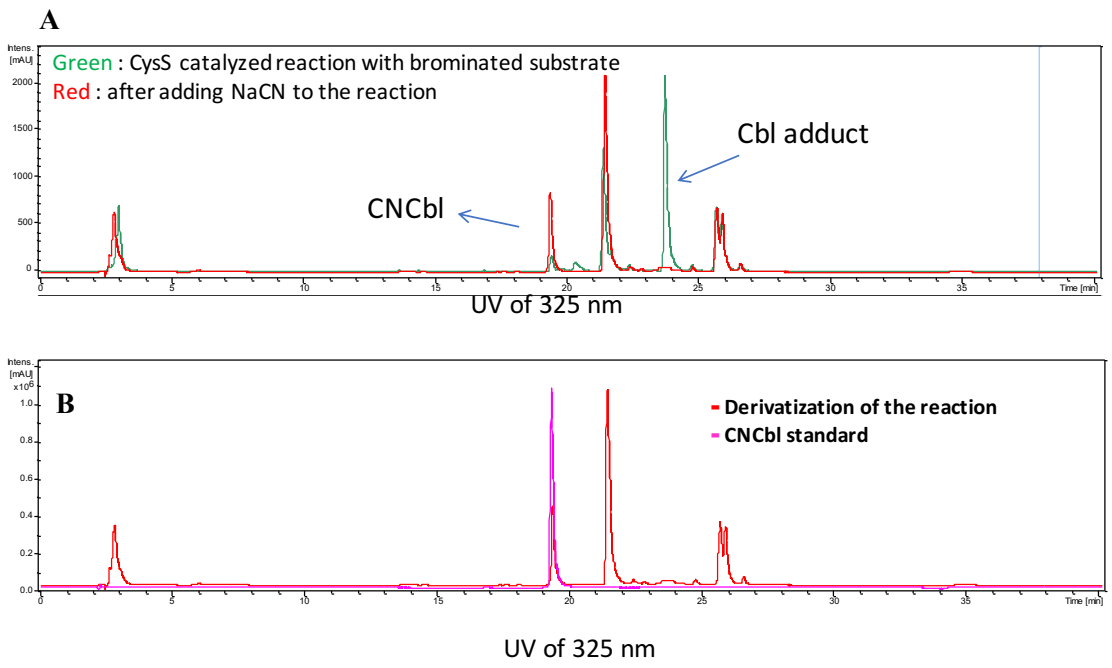


Figure 3.13 HPLC spectrum of derivatization of CysS catalyzed reaction using brominated analog **2-49** as substrate by cyanide

In methionine synthase, Cbl(I) is the active species for the methyl transfer. Since we are not able to detect Cbl(I) by UV during the CysS reaction, it implies that Cbl(I) was transient and strongly coupled with the methyl transfer, which has also been observed for methionine synthase. However, we are able to design an electrophile Br-NAC **2-49** to trap the nucleophile Cbl(I) species in the CysS active site. In CysS, Cbl(I) is also responsible for the regeneration of MeCbl using SAM as a methyl donor.

Here, Cbl(I) is methylated by Br-NAC **2-49** instead of SAM in the native reaction. This suggests that Br-NAC **2-49** is a better substrate for Cbl methylation than SAM, and the ethyl group of **2-49** is closer to cobalt than the methyl group of SAM, or it blocks SAM to methylate Cbl in a certain amount of time. Since a large amount of 5′-dA

was not detected when Br-NAC **2-49** was used as a substrate for CysS, it indicated that radical chemistry was not favored unless the MeCbl was formed. When Cbl substrate crosslinking species was incubated with SAM, CysS and the reducing system, an increased amount of 5′-dA was detected (not as much as the native substrate). This indicated that radical chemistry was not initiated until the Cbl gets methylated, and the substrate is present. However, the binding affinity of the Cbl substrate crosslinking species probably is not as good as the native substrate. That's why following radical mediated methylation step was not observed for the brominated substrate analog. The Cbl substrate crosslinking species closely mimics the conformation when the methyl substrate is methylated by MeCbl to form ethyl group and suggests that the substrate is situated between MeCbl and SAM (Figure 3.14).

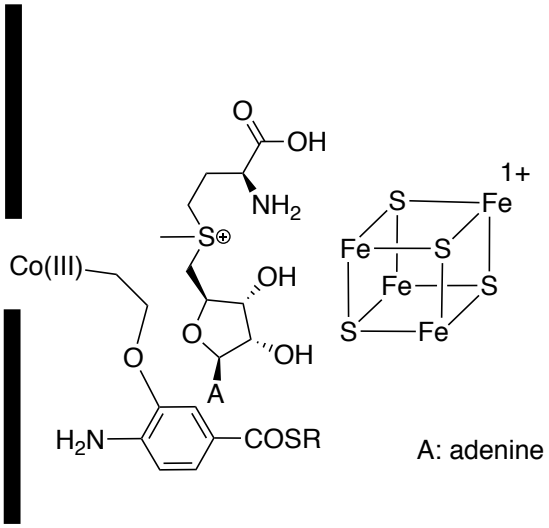


Figure 3.14 Proposed orientation of adduct and cofactors in the CysS active sites

3.2.3. Identification of physiological cobalamin cofactor for CysS

As MeCbl can be regenerated in enzyme active through Cbl(I) species, we are wondering whether MeCbl function as an enzyme cofactor or a substrate for CysS. CD_3Cbl was synthesized and incubated with CysS, SAM, MV, NADPH. Unexpectedly, a majority of the methyl group was deuterium labeled even under single turnover condition (Figure 3.15). Furthermore, when CD_3SAM was used with an excess of CH_3Cbl, 76% of isopropyl product was deuterium labeled (When mixed MeCbl and OHCbl were used with CD_3SAM, the percentage of deuterium incorporation will increase with the increased amount of OHCbl). It was found that OHCbl can be reduced to Cbl(II) more readily by MV, NADPH than MeCbl. These results indicate that CysS prefer the non-methylated Cbl than the methylated Cbl. CysS can perform sequential methylations with a catalytic amount of Cbl, which suggests that MeCbl is regenerated in situ in the enzyme active site. A S_N2 mechanism requires in-line geometry of the methyl group donor- SAM and the methyl group acceptor-Cbl. Br-NAC **2-49** analog can also have a similar methyl transfer reaction, which suggests that the substrate and SAM probably share one binding site in at least one of the conformation. Thus, it's very likely that SAM first methylates Cbl and then the substrate enters into the enzyme active site followed by the hydrogen abstraction, methylation. Then, the methylated substrate left the active site or have a substantial conformational change to allow the methyl group of the SAM align on the top of cobalamin for re-methylation for the next catalytic cycle.

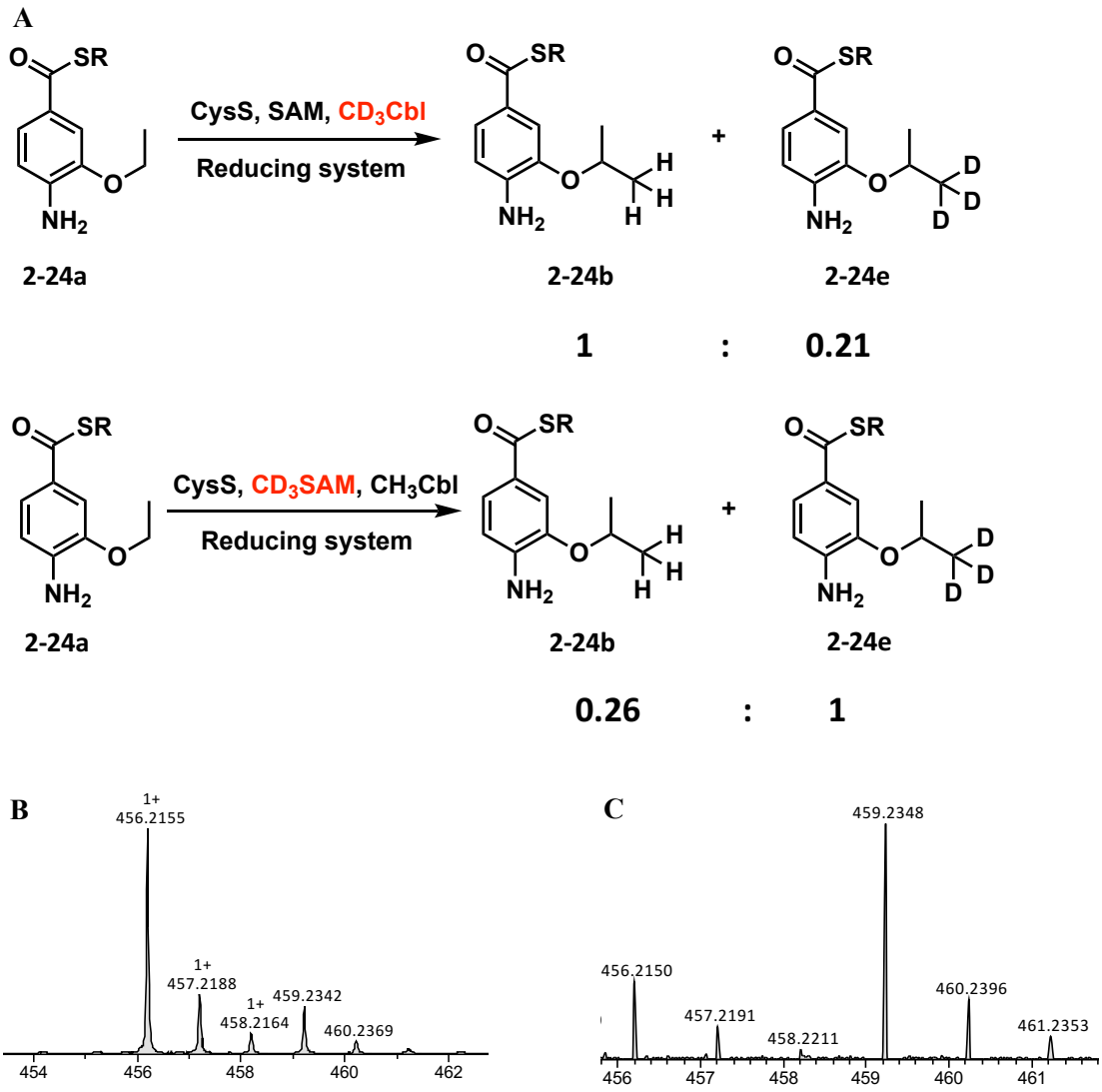


Figure 3.15 CysS catalyzed methylation with (A) CD_3Cbl, SAM (top) or CD_3 SAM, CH_3Cbl (bottom) in the presence of reducing system. (B) MS spectrum of the isopropyl product when CD_3Cbl, SAM was used. (C) MS spectrum of the isopropyl product when CD_3 SAM, CH_3Cbl was used.

3.2.4. Measurement of the competitive and intrinsic isotope effect

The vast majority of radical SAM enzymes are known for generation of 5′dA•, which can abstract a hydrogen atom from the substrate. Isotope effects have been used as a powerful tool to interrogate transition states of carbon hydrogen bond cleavage. CD_3-

pantethinyl substrate **2-53** was synthesized and mixed with its protium substrate **2-23** in 1:1 ratio in order to measure the intermolecular isotope effect. The enzymatic reaction was quenched at different time point (10 mins, 30 mins, 1h, 2.5h, 4h, 5.5h) and then analyzed by LC-MS/MS. The ratio of the methylated protium products **2-24a** and the methylated deuterium product **2-54** was measured and the competitive isotope effect is 17 ± 0.3.

Usually, the kinetic isotope effect can be masked by slower chemical steps which are not isotope-sensitive and by slow product release. In order to obtain the intrinsic isotope effect for hydrogen abstraction, CD_2H-pantetheinyl substrates **2-55** were synthesized. Here, we assumed that the three hydrogens on the methyl group could rotate freely in the enzyme active site and substitutions with deuterium will not have a significant effect. The KIE for CD_2H-pantetheinyl substrates **2-55** was 18 ± 0.6 measured based on the products ratio (Figure 3.16). The large intermolecular isotope effect indicates that the C-H bond breaking happens before or at the first irreversible step in the CysS enzymatic reaction. The small difference between intrinsic isotope effect and the intermolecular isotope effect also suggests that there is no slow step that precedes the hydrogen abstraction step. Furthermore, all isotope effects are larger than the semiclassical limit predicted by transition state theory. Such large kinetic isotope effects are often interpreted as an indication of quantum mechanical tunneling of the hydrogen atom. This kind of substantial isotope effects has also been reported for other radical SAM enzymes in the literature, like lipoyl synthase, biotin synthase (Figure 3.17).[62-63] In the lipoyl cofactor biosynthesis, lipoyl synthase (LipA) catalyzes the insertion of two

sulfur atoms at C6 and C8 of an (N^6-octanoyl)-lysyl residue on a lipoyl carrier protein (LCP) in the final step of lipoyl cofactor biosynthesis. When 8,8,8-trideuterooctanoyl peptide was used as a substrate for LipA, a large isotope effect was observed with only minute quantities of the lipoyl product which the C8 hydrogen is abstracted[64]. Because of this significant isotopic discrimination, Booker's group was able to trap a covalent cross-link intermediate between the peptide and LipA before the second hydrogen abstraction from C8[65].

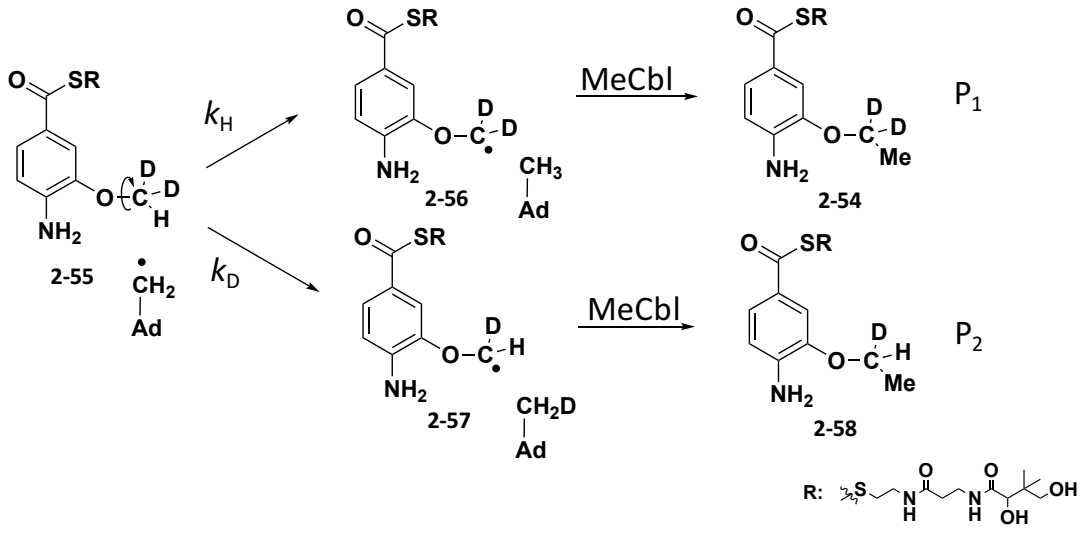


Figure 3.16 The strategy of intrinsic isotope effect measurement

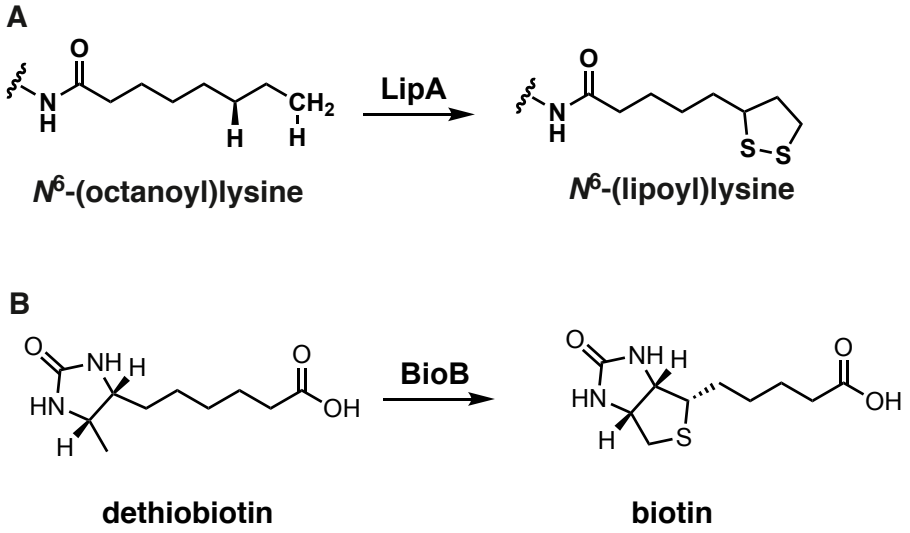


Figure 3.17 (A) LipA in the biosynthesis of lipoic acid; (B) The conversion of dethiobiotin to biotin catalyzed by biotin synthase (BioB).

3.2.5. Efforts to obtain a crystal structure of CysS

Currently, no crystal structure is published for the B_{12} dependent radical SAM methyltransferase. The crystallography studies of CysS is collaborated with Dr. Drennan lab. From the size exclusion column analysis of purified CysS, it's found that when CysS is overexpressed and purified in *E. coli* B12, CysS can easily aggregate to form dimers, trimers until hexamers in solution. In order to solve the heterogeneous problem, different constructs were made for CysS, including cleavable His tag site on both C-terminal and N-terminal of the protein, a SUMO tagged and a C-terminal truncated version of CysS. We also overexpressed and made the peptide carrier protein CysG1 and native substrate of CysS – small molecule modified CysG1 to co-express or incubate with CysS in order to get homogeneous protein for crystallography studies. The best condition was obtained when CysS was co-expressed with a cobalamin uptake and

transport system in *E. coli*, developed by Dr. Booker.[24] The cobalamin content and monomer species of CysS were increased. The crystallography studies are still in progress.

3.3. Conclusion

Based on the above result, we can build a detailed model of how CysS uses Cbl and radical SAM machinery to sequential methylate a methyl group to form ethyl, isopropyl, *sec*-butyl, and *t*- butyl groups (Figure 3.18). One molecule of SAM first binds to the [4Fe–4S] cluster in a conformation that can methylate Cbl. After the Cbl(I) is methylated by SAM via the S_N2 mechanism, S-adenosyl homocysteine(SAH) is released. A second molecule of SAM binds to [4Fe–4S] cluster, because of the steric hindrance of MeCbl, the binding mode of the second molecule of SAM will be different from the first one, which favors the reductive cleavage and also opens up the active site for the binding of the substrate. Once the substrate enters into the active site, [4Fe–4S] cluster will cleave SAM to generate 5′-dA •, which abstracts a hydrogen atom from the methyl group of the substrate **2-23**. The resulting radical **2-25** then undergoes a radical substitution with methylcobalamin to give the ethyl ether **2-24a**. Regeneration of MeCbl from Cbl(II) can be achieved by reduction to Cbl(I) by the $[4Fe{-}4S]^{+1}$ cluster followed by SAM-mediated methylation. Repetition of this sequence results in the successive formation of the isopropyl, *t*-butyl, and *sec*-butyl ethers of **2-23**.

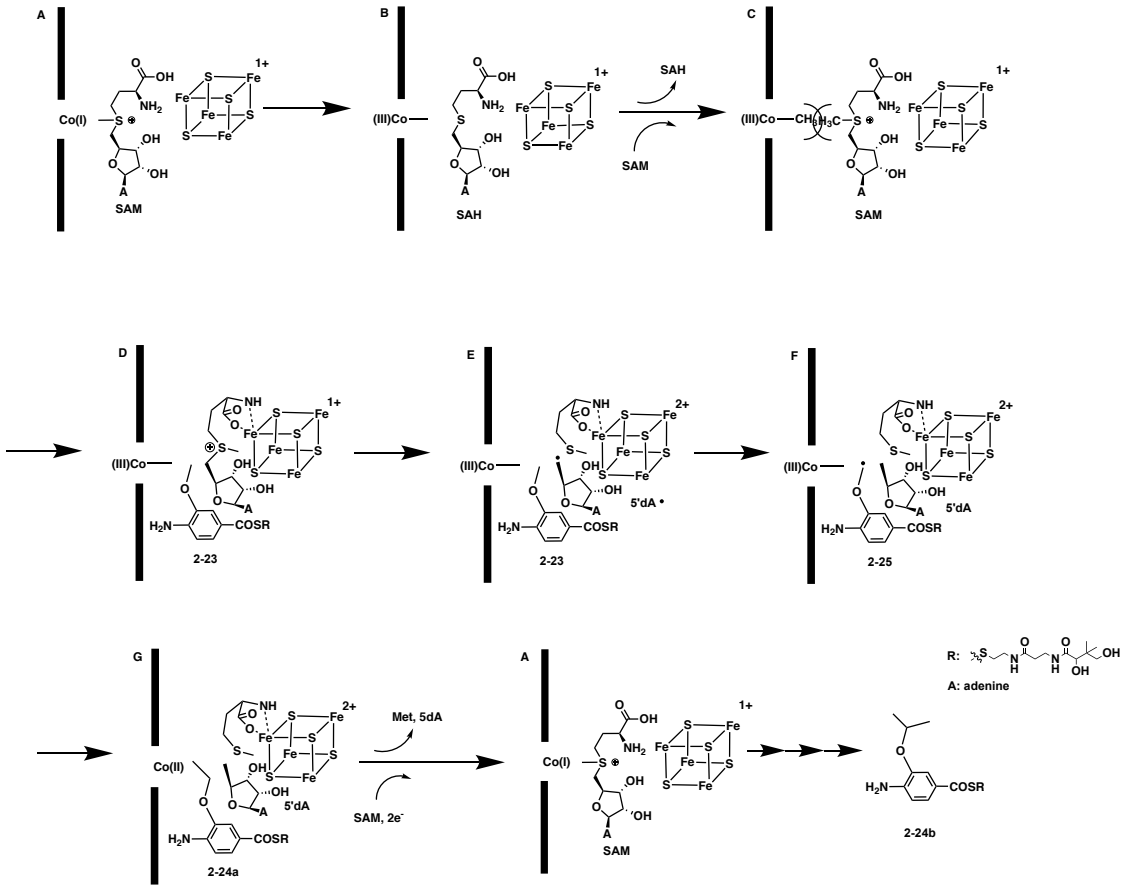


Figure 3.18 Proposed model for CysS catalyzed sequential methylation.

Based on the isotope labeling cofactors experiment and Br-NAC **2-49** probe, we propose that Cbl is first reduced to Cbl(I) and then get methylated by SAM (Figure 3.18A-B). No 5′-dA was detected when Cbl is absent in the reaction. The result of an increased amount of 5′-dA was detected when Cbl substrate crosslinking species was used as the substrate for CysS indicates that radical chemistry was not initiated until the Cbl get methylated, and the substrate is present. The formation of Cbl substrate crosslinking species **2-50** also suggests that the substrate is sandwiched between the Cbl cofactor and the radical SAM machinery (Figure 3. 18D-G), which is different from

GenK. For GenK, hydrogen abstraction is stereospecific and the methylation occurs with retention of configuration at C6′ of the substrate, which indicates MeCbl donor likely sits adjacent to the radical SAM machinery in the GenK active site. For Fom3, GenD1, and MoeK5, the methylations occur with inversion of configuration at the carbon center. These suggest that MeCbl and the radical SAM machinery are probably situated on the opposite sides of the center to be methylated once bound in the active site. Here, the Cbl substrate crosslinking species generated in the CysS active site instead of MeCbl further reassembles the geometry of substrates, cofactors in the later cases.

Based on our experimental results and OxsB crystal structure, SAM is proposed to have one binding site but two conformations for methyl transfer and radical initiation in CysS active sites (Figure 3. 18C-D). A radical clock probe is used to trap the substrate radical (Figure 3. 18F) and estimate the radical substitution rate of a methyl group from cobalamin for the first time. From the large competitive isotope effect, we know that CysS is a distributive rather than a processive enzyme, with the sequential reaction allowing free dissociation of the ethyl and isopropyl products.

To conclude, we are able to develop different substrate analogs and isotope labeling compounds to trap important intermediates and build a model for the sequential methylation of CysS. It is known that the C-Co in MeCbl is considered weak (BDE: 129.7kJ/mol). Thus, nature uses MeCbl as a catalyst to perform radical substitution for unusual methylation on sp^3-hybridized carbon on substrate radical. Here, CysS is able to take advantage of these two complicated pieces of machinery- radical SAM and cobalamin cofactors to perform sequential methylations to form *t*-butyl group by radical

chemistry. The detail mechanistic study of CysS will help to unravel other methyltransferases in this large B_{12} radical SAM superfamily.

3.4. Experimental methods

3.4.1. Quantification of enzymatic products in the CysS catalyzed cyclopropyl substrate analog reaction

A typical enzymatic reaction was performed in 100 mM phosphate buffer, pH 7.5 containing 68 μM of CysS with 0.5 mM MeCbl, 0.5 mM SAM, 0.5 mM substrate with reducing systems (0.8 mM methyl viologen, 65 mM NADPH). The reaction was incubated at room temperature in an anaerobic chamber containing 95% nitrogen and 5% hydrogen. For LC-MS analysis the enzyme was removed by ultrafiltration using a 10kDa cut-off filter (VWR). The enzymatic assay was split into equal amounts. One of them was used to quantified the methylated cyclopropyl product, the other half was used to quantify the ring opening derived product aminophenol compound by adding TFA (final concentration 0.1%). Calibration curves were constructed by measuring the peak intensity of synthetic standard **2-46** and **2-43** with known concentration using LC-MS. Then the calibration curves were used to measure the concentration of **2-46** and **2-43** in the enzymatic reaction. The ratio of the methylted products and the ring opening products is 12.0. If we used the literature reported value (2×10^7 s^{-1}) to estimate the value of the ring opening rate constant of the cyclopropyl substrate (k_{ring}) in CysS reaction, we can calculate the rate constant of the methyl group transfer from MeCbl to substrate radical (k_{me}) as 2.4×10^8 s^{-1}.

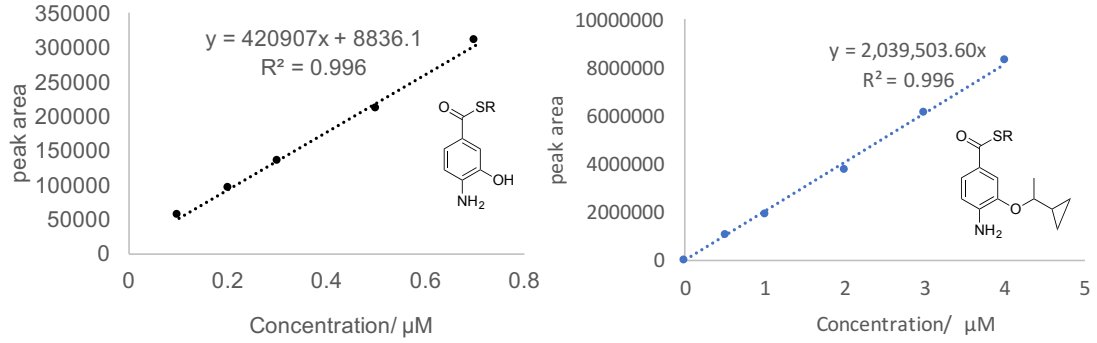


Figure 3.19 calibration curves for the ring opening products and the methylated product

3.4.2. Model reaction for generating the Cbl crosslinking species 2-50

The preparation of Cbl(I) species is followed the protocol in the literature. A solution of hydroxocobalamin (50 mg) and cobalt nitrate (1 mg) in water (10 ml) is deoxygenated with a stream of argon. After 10 min a deoxygenated solution of sodium borohydride (15 mg) in water (1 ml) is added to the red cobalamin solution, which immediately turns brown and then blue-green, signifying that the reduction is complete. The reaction was also monitored by UV in anaerobic environment.

3.4.3. Generation and derivatization of the Cbl crosslinking species in CysS catalyzed reaction

90 µM of CysS was incubated with 1 mM MeCbl, 1 mM SAM, 1.2 mM Br-NAC **2-49** with reducing systems (1.2 mM methyl viologen, 3.5 mM NADPH) at room temperature in an anaerobic chamber in dark for 12-15 hours.

3.4.4. Measurement of the competitive isotope effect of CysS

80 µM of CysS was incubated with 1 mM MeCbl, 1 mM SAM, 1 mM of 87pantCH_3 **2-23** and 87pantCD_3 **2-53** with reducing systems (1 mM methyl viologen, 3

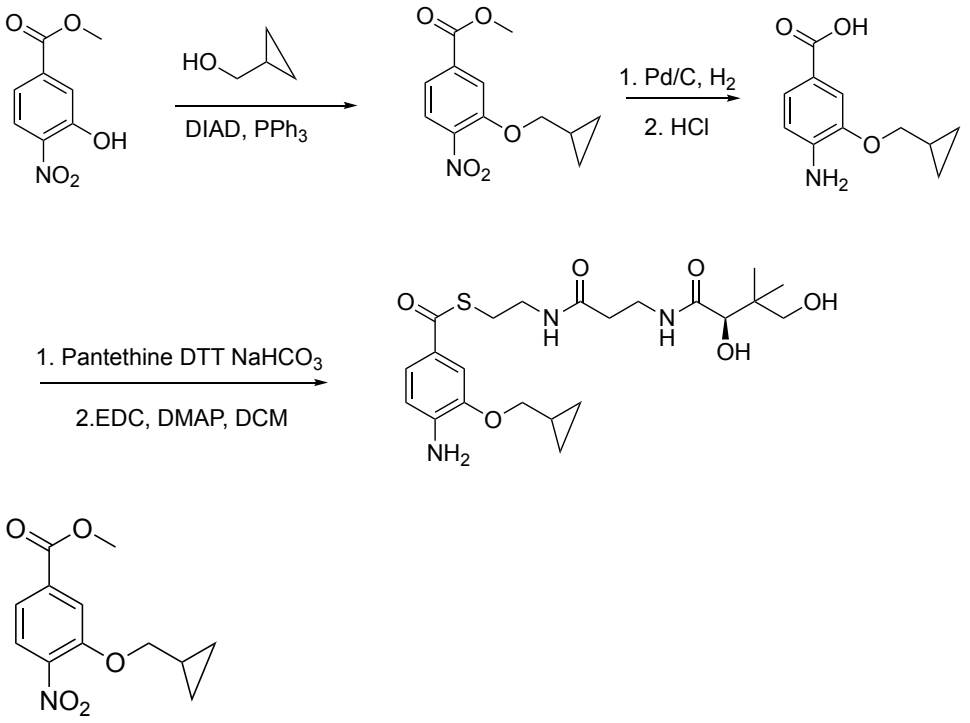
mM NADPH) at room temperature in an anaerobic chamber. The reaction was quenched at different time point (10 mins, 30 mins, 1h, 2.5h, 4h, 5.5h) by ultrafiltration using a 10kDa cut-off filter. The reaction mixtures were analyzed by LC-MS/MS and the ratio of the methylated protium products **2-24a** and the methylated deuterium **2-54** product was measured. Here, 10 min time point was used to obtain the competitive isotope effect, because after 10mins the dimethylated products can be detected by LC-MS/MS.

3.4.5. Measurement of the intrinsic isotope effect of CysS

20 μM of CysS was incubated with 270 μM MeCbl, 540 μM SAM, 540 μM of 87pantCD_2H with reducing systems (1 mM methyl viologen, 3 mM NADPH) at room temperature in an anaerobic chamber. The reaction was quenched at different time point (30 mins, 1h, 2h, 3h, 4h) by ultrafiltration using a 10kDa cut-off filter. The reaction mixtures were analyzed by LC-MS/MS and the ratio of the CDHMe products P1 and CD2Me products P2 was measured. Under this condition, the dimethylated products were not formed in 4 hours. After the natural abundance correction of stable isotopes, the intrinsic isotope effect is 18 ± 0.6 using below equation:

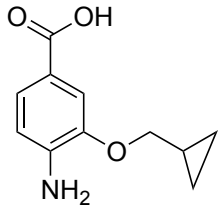
$$r = \frac{k_H}{\frac{1}{2}k_D} = \frac{[P_1]}{\frac{1}{2}[P_2]}$$

3.4.6. Synthesis of (R)-S-pantetheinyl-4-amino-3-(cyclopropylmethoxy)-benzothioate



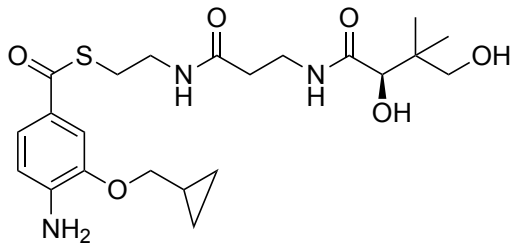
methyl 3-(cyclopropylmethoxy)-4-nitrobenzoate (2-61)

Diisopropyl azodicarboxylate (DIAD, 1.5 mmol) was added to the THF solution of triphenylphosphine (PPh_3, 2.0 mmol) at 0° C. The reaction mixture was stirred for half an hour. Then to it cyclopropylmethanol (1.0 mmol) and methyl 3-hydroxy-4-nitrobenzoate (1.0 mmol) were added one by one. It was then stirred at rt. in dark for overnight. Then it was concentrated and purified by column chromatography (Si-gel, 25% ethyl acetate-hexane) to afforded **2-61**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.77 (d, J = 8.3 Hz, 1H), 7.69 (m, 1H), 7.67 – 7.64 (m, 1H), 4.02 (d, J = 6.8 Hz, 2H), 3.93 (s, 3H), 1.28 (m, 1H), 0.64 (m, 2H), 0.38 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.26, 134.66, 132.10, 128.60, 125.13, 121.32, 116.09, 74.62, 52.76, 9.85, 3.30.



4-amino-3-(cyclopropylmethoxy)benzoic acid (2-62)

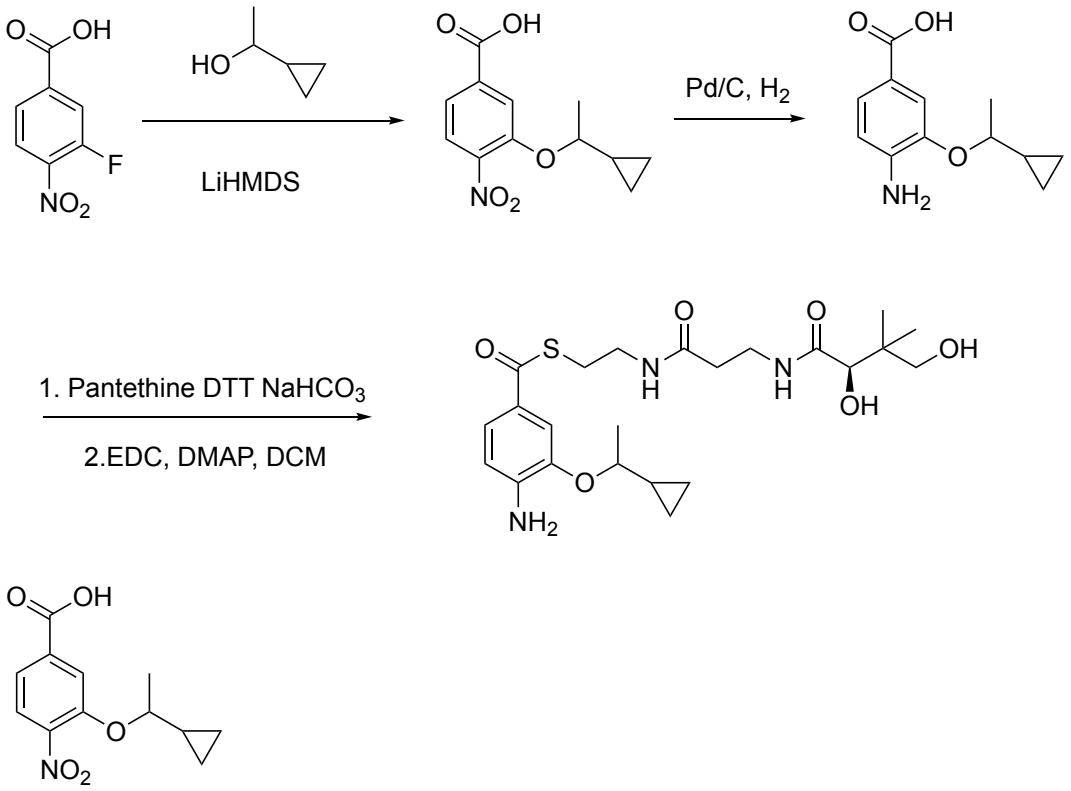
Using 10% Pd/C as catalyst, 50 mL of methyl methyl 3-(cyclopropylmethoxy)-4-nitrobenzoate **2-62** (1.0 mmol) in methanol was pumped through the H-Cube®. The pressure of the system was set to full hydrogen mode, and the temperature to 35 °C. The flow rate was 0.8 mL/min. The fraction was analyzed using TLC, which showed complete conversion to the product, and the solvent was reduced to dryness, affording the 4-aminobenzoic acids. The resulting methyl 4-amino-benzoate (1.0 mmol) was suspended in 5M HCl (15 ml) and the mixture was heated at 60 °C. The reaction was monitored by TLC till completed. The reaction mixture was concentrated under vacuum and purified by column chromatography (silica gel, 40:1 CH_2Cl_2: MeOH) to give **2-62**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.60 (dd, J = 8.2, 1.5 Hz, 1H), 7.44 (d, J = 1.4 Hz, 1H), 6.66 (d, J = 8.2 Hz, 1H), 3.88 (d, J = 6.9 Hz, 2H), 1.30 – 1.25 (m, 1H), 0.62 (m, 2H), 0.34 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.96, 145.42, 142.27, 125.06, 118.32, 113.08, 104.99, 77.33, 53.39, 10.27, 3.18.



(R)-S-pantetheinyl-4-amino-3-(cyclopropylmethoxy)-benzothioate (2-41)

D-pantethine (0.2 mmol) was reduced to D-pantetheine by addition of DTT (0.2 mmol) and stirring for 1 h at room temperature. Water was removed by vacuum to leave a clear, viscous liquid. To a stirred solution of 4-amino-3-(cyclopropylmethoxy)benzoic acid **2-41** (0.4 mmol) in 3 mL of dichloromethane was added 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC, 0.4 mmol), followed by 4-dimethylaminopyridine (DMAP, 0.05 mmol) at room temperature. The resulting mixture was added to the pantetheine solution. The reaction was stirred at room temperature for overenight. All solvent was removed by vacuum and was purified by column chromatography (silica gel, 40:1 CH_2Cl_2: MeOH) to give **2-41**. ^{1}H NMR (400 MHz, MeOD) δ 8.19 (s, 1H), 7.94 (s, 1H), 7.50 (dd, J = 8.3, 1.9 Hz, 1H), 7.35 (d, J = 1.8 Hz, 1H), 6.70 (d, J = 8.3 Hz, 1H), 3.94 – 3.87 (m, 3H), 3.44 (m, 6H), 3.16 (t, J = 6.6 Hz, 2H), 2.43 (t, J = 6.6 Hz, 2H), 1.31 (m, 1H), 0.93 (s, 6H), 0.70 – 0.58 (m, 2H), 0.44 – 0.35 (m, 2H). ^{13}C NMR (101 MHz, MeOH) δ 189.61, 174.70, 172.60, 145.29, 144.14, 125.32, 122.54, 112.20, 109.78, 76.00, 73.08, 69.03, 39.25, 38.96, 35.10, 34.98, 27.53, 19.91, 19.56, 9.79, 2.21.

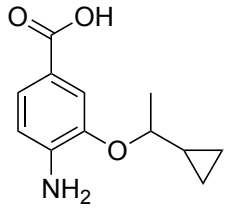
3.4.7. Synthesis of (R)-S-pantetheinyl-4-amino-3-(1-cyclopropylethoxy)-benzothioate



3-(1-cyclopropylethoxy)-4-nitrobenzoic acid (2-63)

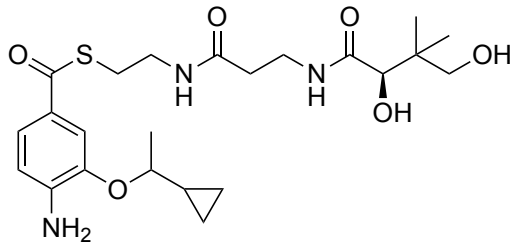
3-fluoro-4-nitrobenzoic acid (1.0 mmol) and 1-cyclopropylethan-1-ol (1.1 mmol) are added to lithium bis(trimethylsilyl)amide (LiHMDS) (1M in THF, 1.1 ml) at 0 °C. KHMDS. 3.5 mL of LiHMDS solution was then added dropwise, resulting in a color change from light yellow to orange. This solution is stirred from 0 °C to room temperature over 2 h and then diluted with CH_2Cl_2 and washed once with saturated aqueous NH_4Cl. The aqueous layer is back-extracted once with CH_2Cl_2. The organic layers are combined, dried over MgSO4, concentrated, and purified by column chromatography (silica gel, 20:1 CH_2Cl_2: MeOH) to give **2-63**. ^{1}H NMR (400 MHz,

MeOD) δ 7.87 – 7.75 (m, 2H), 7.70 (m, 1H), 4.20 (dt, J = 12.5, 6.2 Hz, 1H), 1.44 (d, J = 6.1 Hz, 3H), 1.14 (m, 1H), 0.57 (m, 2H), 0.36 (m, 2H). ^{13}C NMR (101 MHz, MeOD) δ 165.73, 149.90, 143.68, 134.46, 123.69, 120.82, 117.28, 79.95, 17.92, 15.42, 2.21, 0.24.



4-amino-3-(1-cyclopropylethoxy)benzoic acid (2-64)

Compound **2-64** was prepared by the above procedure for methyl 4-amino-3-(cyclopropylmethoxy)benzoic acid **2-63**. ^{1}H NMR (400 MHz, MeOD) δ 7.55 – 7.42 (m, 2H), 6.95 (m, 1H), 3.93 – 3.81 (m, 1H), 1.39 (dd, J = 6.1, 2.0 Hz, 3H), 1.14 (dt, J = 8.0, 5.0 Hz, 1H), 0.58 – 0.48 (m, 2H), 0.41 – 0.24 (m, 2H). ^{13}C NMR (101 MHz, MeOD) δ 146.30, 144.19, 124.46, 123.80, 115.67, 113.07, 112.48, 79.38, 18.74, 16.41, 2.94, 0.88.

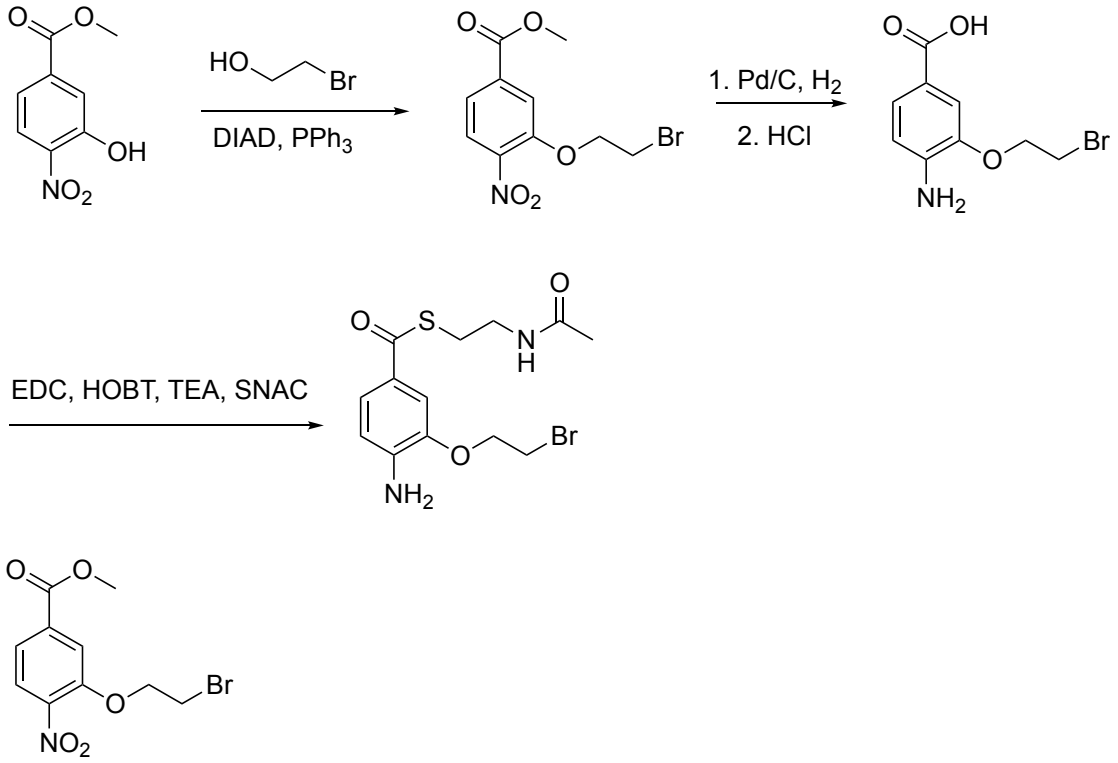


(R)-S-pantetheinyl-4-amino-3-(1-cyclopropylethoxy)-benzothioate (2-43)

Compound **2-43** was prepared by the above procedure for (R)-S-pantetheinyl-4-amino-3-(cyclopropylmethoxy)-benzothioate **2-41**. ^{1}H NMR (400 MHz, MeOD) δ 7.49 (dd, J = 8.3, 1.9 Hz, 1H), 7.39 (d, J = 1.8 Hz, 1H), 6.71 (d, J = 8.3 Hz, 1H), 3.92 – 3.84 (m, 2H), 3.55 – 3.38 (m, 6H), 3.16 (t, J = 6.6 Hz, 2H), 2.43 (t, J = 6.6 Hz, 2H), 1.41 (d, J = 6.1 Hz, 3H), 1.21 – 1.11 (m, 1H), 0.93 (s, 6H), 0.56 (d, J = 8.2 Hz, 2H), 0.43 – 0.28

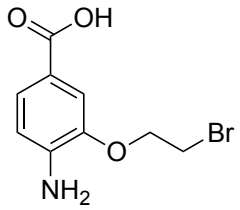
(m, 2H). ^{13}C NMR (101 MHz, MeOD) δ 189.57, 174.62, 172.51, 145.28, 144.17, 125.27, 122.65, 112.81, 112.59, 79.41, 75.99, 69.02, 39.11, 38.96, 35.04, 34.96, 27.51, 19.90, 19.55, 18.70, 16.36, 3.01, 0.86.

3.4.8. Synthesis of S-(2-acetamidoethyl) 4-amino-3-(2-bromoethoxy)-benzothioate



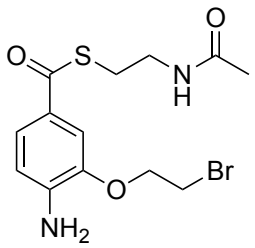
methyl 3-(2-bromoethoxy)-4-nitrobenzoate (2-65)

Compound **2-65** was prepared by the above procedure for methyl 3-(cyclopropylmethoxy)-4-nitrobenzoate using 2-bromoethan-1-ol. ^{1}H NMR (400 MHz, MeOD) δ 7.87 – 7.80 (m, 2H), 7.73 (m, 1H), 4.57 (t, J = 11.4 Hz, 2H), 3.94 (s, 3H), 3.75 (t, J = 11.4 Hz, 2H). ^{13}C NMR (101 MHz, MeOD) δ 172.95, 166.49, 152.06, 136.03, 126.11, 123.26, 117.12, 71.24, 53.27, 29.45.



4-amino-3-(2-bromoethoxy)benzoic acid (2-66)

Compound **2-66** was prepared by the above procedure for methyl 4-amino-3-(cyclopropylmethoxy)benzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.76 (m, 2H), 7.52 (m, 1H), 4.3 (t, J = 11.4 Hz, 2H), 3.81 (t, J = 11.4 Hz, 2H).^{13}C NMR (101 MHz, MeOD) δ 168.03, 152.58, 134.00, 125.04, 124.31, 124.13, 115.11, 53.04, 29.63.

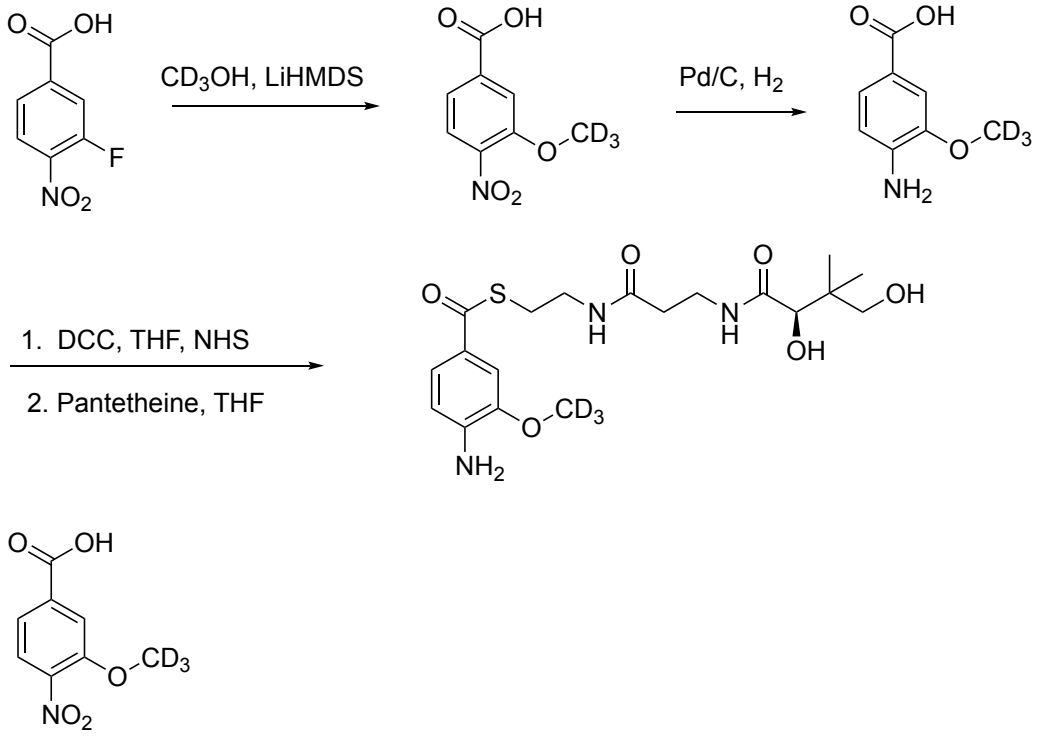


S-(2-acetamidoethyl) 4-amino-3-(2-bromoethoxy)-benzothioate (2-49)

To a solution of triethylamine (TEA, 1.0 mmol) in dichloromethane (3 mL) was added 4-amino-3-(2-bromoethoxy)benzoic acid **2-66** (0.5 mmol), (3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDC, 0.5 mmol), 1-hydroxybenzotriazole (0.5 mmol) and N-acetylcysteamine (SNAC, 0.5mmol). The reaction mixture was stirred overnight under argon. The organic layer was washed with saturated $NaHCO_3$ solution, 0.1 N HCl solution and brine. It was then dried over anhydrous sodium sulfate and concentrated under vacuum. Purification of the crude mixture by flash chromatography (silica gel, 1:1 to 2:1 of hexanes:EtOAc) afforded **2-49**. ^{1}H NMR (400 MHz, MeOD) δ 7.50 (dd, J = 8.3, 1.9 Hz, 1H), 7.37 (d, J = 1.9 Hz,

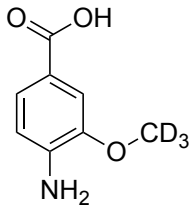
1H), 6.70 (d, J = 8.3 Hz, 1H), 4.34 (t, J = 11.4 Hz, 2H), 3.74 (t, J = 11.4 Hz, 2H), 3.38 (t, J = 6.7 Hz, 2H), 3.13 (t, J = 6.7 Hz, 2H), 1.92 (s, 3H). ^{13}C NMR (101 MHz, MeOD) δ 190.83, 173.45, 145.84, 145.75, 126.68, 124.63, 113.91, 111.95, 70.15, 40.54, 30.62, 28.94, 22.51.

3.4.9. Synthesis of (R)-S-pantetheinyl-4-amino-3-(methoxy- d_3)-benzothioate



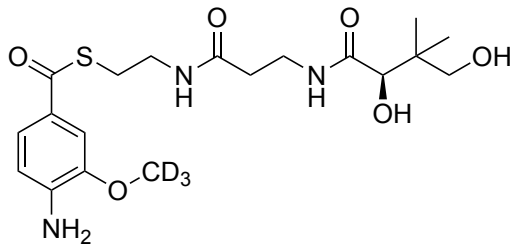
3-(methoxy-d_3)-4-nitrobenzoic acid (2-67)

Compound **2-67** was prepared by the above procedure for methyl 4-amino-3-(cyclopropylmethoxy)benzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.91 – 7.77 (m, 2H), 7.70 (dd, J = 8.3, 1.5 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 167.68, 153.46, 144.08, 136.99, 125.93, 122.67, 115.80, 73.72.



4-amino-3-(methoxy- d_3) benzoic acid (2-68)

Compound **2-68** was prepared by the above procedure for methyl 4-amino-3-(cyclopropylmethoxy)benzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.49 – 7.37 (m, 2H), 6.68 (d, J = 8.2 Hz, 1H). ^{13}C NMR (101 MHz, MeOD) δ 170.65, 147.41, 144.05, 125.58, 119.55, 113.85, 112.29.

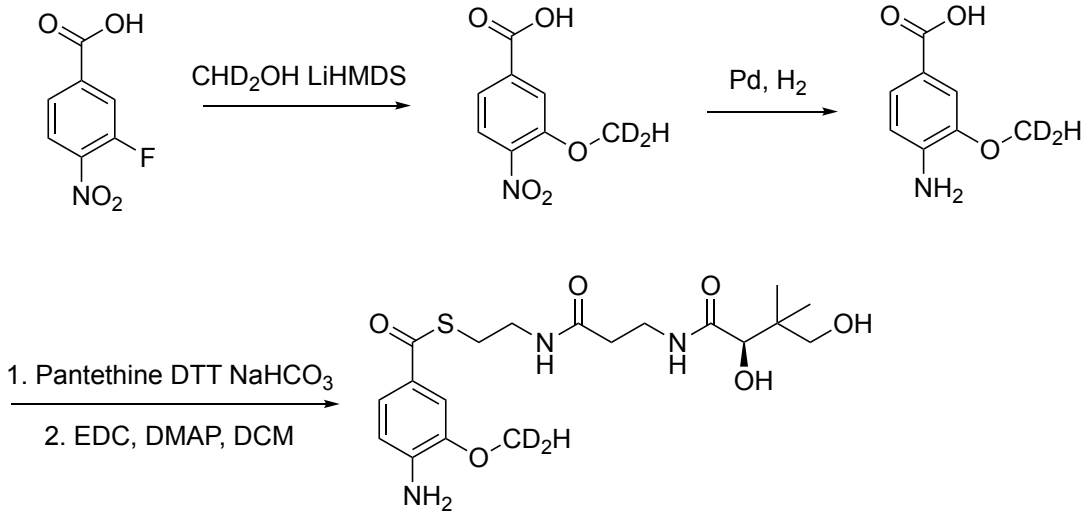


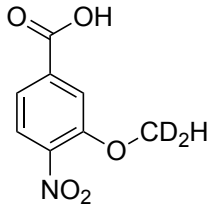
(R)-S-pantetheinyl-4-amino-3-(methoxy- d_3)-benzothioate (2-53)

To a stirred solution of 4-amino-3-(methoxy- d_3) benzoic acid **2-68** (0.5 mmol) in 10 mL of dry THF was added N-hydroxysuccinimide (NHS) (0.5 mmol), followed by N, N'-Dicyclohexylcarbodiimide (DCC) (0.5 mmol) at room temperature. The resulting mixture was stirred overnight at room temperature. The white precipitate was filtered and then washed with EtOAc. The filtrate was concentrated under reduced pressure. The residue was triturated with CH_2Cl_2 to give the NHS ether of **2-68**. Pantethine (0.075 mmol), sodium bicarbonate (1.9 mmol), and 1,4-dithiothreitol (DTT) (0.08 mmol) were dissolved in 3 mL of water, and left to stand for 10 minutes to allow reduction of the pantethine to pantetheine to take place. The NHS ester (0.13 mmol) was dissolved in 6

mL of tetrahydrofuran (THF) and added to the 3 mL water mixture. The reaction mixture was stirred overnight and then the solvent was evaporated under reduced pressure. The crude product was then dissolved in water and the insoluble residue was removed by filtration. The resulting solution was lyophilized and purified by flash chromatography (silica gel, 20:1 to 10:1 CH_2Cl_2: MeOH). The fraction was further purified by Prep HPLC using a SPLC-18DB column (10 x 250 mm, 5 µm, Supelco) at a flow rate of 2 mL/min on an Agilent 100 HPLC system with a quaternary pump and an automatic injector. ^{1}H NMR (400 MHz, MeOD) δ 7.48 (dd, J = 8.3, 1.9 Hz, 1H), 7.35 (d, J = 1.9 Hz, 1H), 6.66 (d, J = 8.3 Hz, 1H), 3.88 (s, 1H), 3.59 – 3.33 (m, 6H), 3.14 (t, J = 6.7 Hz, 2H), 2.41 (t, J = 6.6 Hz, 2H), 0.91 (d, J = 1.4 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 191.00, 176.04, 173.92, 147.47, 145.35, 126.70, 123.97, 113.41, 109.57, 77.40, 70.43, 40.54, 40.37, 36.45, 36.38, 28.91, 21.30, 20.95.

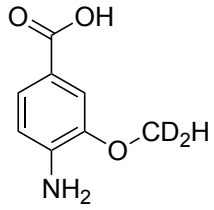
3.4.10. Synthesis of (R)-S-pantetheinyl-4-amino-3-(methoxy- *d_2*)-benzothioate





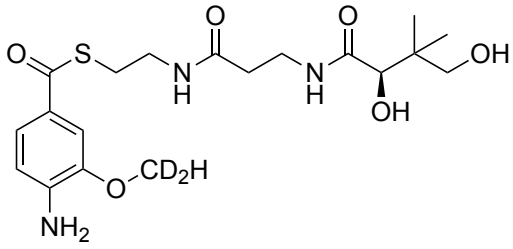
3-(methoxy-d_2)-4-nitrobenzoic acid (2-69)

Compound **2-69** was prepared by the above procedure for methyl 4-amino-3-(cyclopropylmethoxy)benzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.86 – 7.79 (m, 2H), 7.70 (dd, J = 8.3, 1.4 Hz, 1H), 3.96 (s, 1H).



4-amino-3-(methoxy-d_2) benzoic acid (2-70)

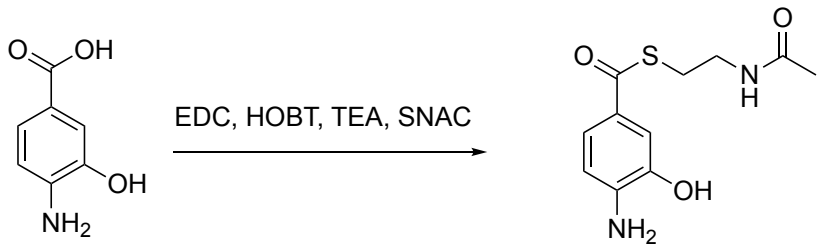
Compound **2-70** was prepared by the above procedure for methyl 4-amino-3-(cyclopropylmethoxy)benzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.50 – 7.35 (m, 2H), 6.68 (d, J = 8.2 Hz, 1H), 3.82 (s, 1H). ^{13}C NMR (101 MHz, MeOD) δ 170.90, 147.44, 143.97, 125.57, 119.88, 113.89, 112.33, 55.69, 55.47, 55.25.



(R)-S-pantetheinyl-4-amino-3-(methoxy-d_2)-benzothioate (2-55)

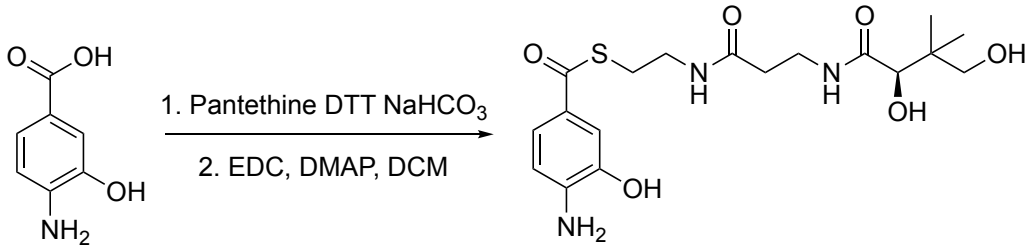
Compound **2-55** was prepared by the above procedure for (R)-S-pantetheinyl-4-amino-3-(cyclopropylmethoxy)-benzothioate using 4-amino-3-(methoxy- d_2) benzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.51 (dd, J = 8.3, 1.9 Hz, 1H), 7.38 (d, J = 1.9 Hz, 1H), 6.69 (d, J = 8.3 Hz, 1H), 3.91 (s, 1H), 3.88 – 3.84 (m, 1H), 3.57 – 3.38 (m, 6H), 3.17 (t, J = 6.7 Hz, 2H), 2.44 (t, J = 6.7 Hz, 2H), 0.94 (d, J = 1.4 Hz, 6H). ^{13}C NMR (101 MHz, MeOD) δ 189.61, 174.63, 172.51, 146.06, 143.93, 125.29, 122.57, 112.01, 108.17, 75.99, 69.02, 54.36, 54.14, 53.92, 39.13, 38.96, 35.04, 34.97, 27.51, 19.91, 19.54.

3.4.11. Synthesis of S-(2-acetamidoethyl) 4-amino-3-hydroxybenzothioate (2-52)



Compound **2-52** was prepared by the above procedure for S-(2-acetamidoethyl) 4-amino-3-(2-bromoethoxy)-benzothioate using 4-amino-3-hydroxybenzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.72 – 7.63 (m, 1H), 7.63 – 7.48 (m, 2H), 3.46 (t, J = 6.7 Hz, 2H), 2.81 (t, J = 6.7 Hz, 2H), 1.93 (s, 3H).

3.4.12. Synthesis of (R)-S-pantetheinyl-4-amino-3-hydroxybenzothioate (2-46)



Compound **2-46** was prepared by the above procedure for (R)-S-pantetheinyl-4-amino-3-(cyclopropylmethoxy)-benzothioate using 4-amino-3-hydroxybenzoic acid. ^{1}H NMR (400 MHz, MeOD) δ 7.40 (dd, J = 8.3, 2.0 Hz, 1H), 7.31 (d, J = 2.0 Hz, 1H), 6.67 (d, J = 8.3 Hz, 1H), 3.91 (s, 1H), 3.51 – 3.38 (m, 6H), 3.14 (t, J = 6.7 Hz, 2H), 2.43 (t, J = 6.7 Hz, 2H), 0.94 (s, 7H). ^{13}C NMR (101 MHz, MeOD) δ 189.67, 174.63, 172.51, 143.42, 143.02, 125.71, 121.23, 112.69, 112.22, 75.98, 69.01, 39.14, 38.95, 35.02, 34.97, 27.41, 19.90, 19.52.

4. TRAPPING RADICAL INTERMEDIATES IN AN ANEROBIC HEME DEGRADATION PATHWAY

4.1. Introduction

Iron is known to play a crucial role in the living system. Although iron is the fourth most common element in the Earth's crust, it is not readily bioavailable at neutral pH due to oxidation to the insoluble ferric oxide. For the soluble reduced Fe^{2+}, it can catalyze the generation of toxic free radicals through Fenton chemistry. These reactive oxygen species (ROS) can react with membrane lipids, proteins, and nucleic acids, which cause irreversible cell damage.[66] Thus, the free iron concentration has to remain very low. Mammals employ heme, iron-binding proteins (transferrin, lactoferrin) to reduce the levels of free extracellular iron to around 10^{-18} M, which is orders of magnitude below the 10^{-6} M required for bacterial growth.[67] The majority of iron within the host, approximately 95%, is sequestered in the form of heme, primarily as hemoglobin.[68] However, free heme is considered toxic to the cells.[69]

In order to compete for the host's iron during infection, pathogens have developed sophisticated mechanisms to utilize heme as a source of iron.[67] Bacteria employ surface proteins to seize heme from hemoproteins and then shuttle heme across the outer-membrane/cell wall onto specific transporters for import across the cytoplasmic membrane.[70] Inside the bacterial cell, heme is incorporated directly into proteins or it is further degraded to release the iron.[71] Oxidative degradation by heme oxygenases (HOs) is the most common mechanism used to liberate iron from heme.

In most of the cases, heme oxygenase catalyzes oxidative cleavage of heme.[72] Heme oxygenase requires three molecules of oxygen and seven electrons to regiospecifically cleave heme at the α-meso-carbon to yield biliverdin IXα via α-meso-hydroxyheme and verdoheme.[73-75]

However, it's known that pathogens can colonize the nonsterile region of the lower intestines, which is an anaerobic environment. For example, *Vibrio cholerae* and *Escherichia coli* O157:H7 are hemolytic enteric pathogens, and a bona fide heme oxygenase has not been characterized for them. Both organisms can use heme as an iron source and contain a heme uptake operon that is up-regulated during iron duress.[76-77]

The heme uptake operon includes a heme uptake and transport machinery, and three additional genes, *chuw*, *chux*, and *chuy* (Figure 4.1). ChuW is assigned as a radical SAM enzyme and has high sequence similarity with coproporphyrinogen III oxidase (HemN), which is a characteristic feature of Class C radical SAM methyltransferases. Class C RSMTs are proposed to use two molecules of SAM to methylate sp^2-hybridized carbon centers.[19]

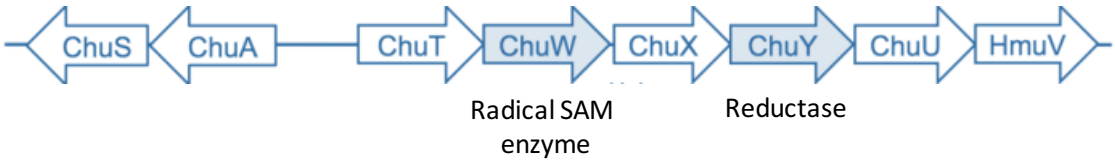


Figure 4.1 The heme uptake operon in *Escherichia coli* O157:H7

In 2016, Dr. Lanzilotta demonstrates that ChuW is able to catalyze a radical-mediated reaction to break the protoporphyrin ring of heme to release iron and form a tetrapyrrole product "anaerobilin" anaerobically (Figure 4.2).[78] 5′-dA and SAH were produced at 1:1 stoichiometry in the presence of heme, ChuW and flavodoxin, NADPH/flavodoxin reductase system. When [^{13}C-methyl]- SAM was used in the assay, the isolated product exhibited an enrichment of ^{13}C. This suggests that ChuW also catalyzes the transfer of a methyl group from the SAM moiety to the product. Furthermore, the anaerobilin can be used as a substrate by anaerobilin reductase ChuY in the same operon. Nevertheless, the deuteroheme-derived tetrapyrrole product deuteroanaerobilin (DAB) is not stable under the light condition. Exposure of DAB to sunlight for 5 min resulted in a significant change in the UV-visible spectrum. The light sensitivity of anaerobilin and DAB has prohibited the accumulation of product for NMR characterization.

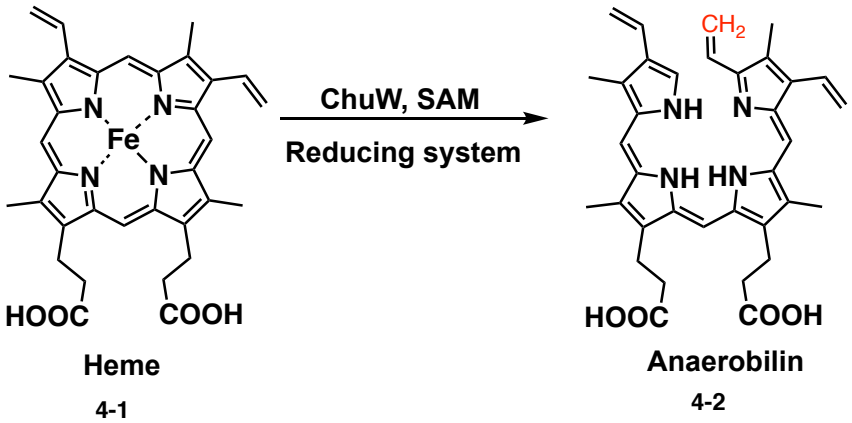


Figure 4.2 ChuW-catalyzed degradation of heme anaerobically

Here, we try to design substrate analogs to trap the possible racial intermediates generated on the tetrapyrrole ring during the ChuW catalyzed reaction in order to study the mechanism.

4.2. Result and discussion

4.2.1. Reconstitution of ChuW catalyzed heme degradation

The chuW gene was obtained from Dr. Lanzilotta and was cloned into a pTHT vector. When heme was used as a substrate for ChuW with *E. coli* flavodoxin (EcFldA)/ferredoxin (flavodoxin): $NADP^+$ oxidoreductase (EcFpr)/ NADPH system as the reducing system, a new species (455nm, 800nm) was formed similarly as the reported anaerobilin with heme degraded over time based on the UV spectrum (Figure 4.3). The proposed anaerobilin was not stable and degraded under light conditions for several minutes.

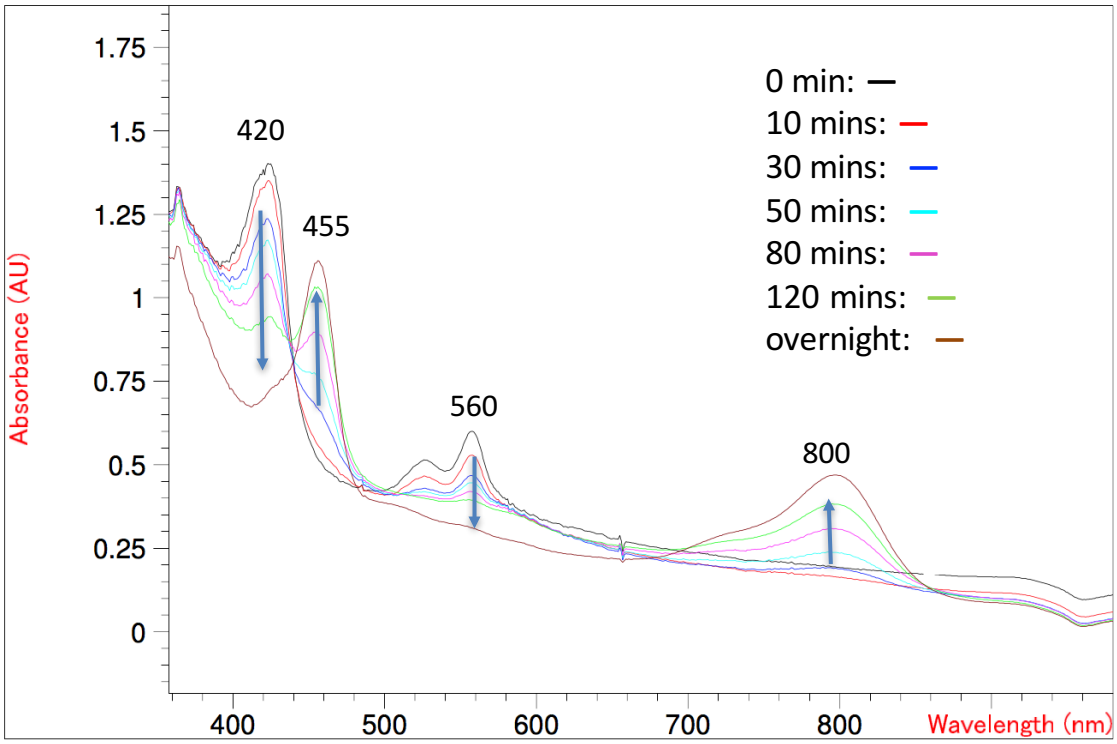


Figure 4.3 UV spectrum of the time course of the ChuW catalyzed heme degradation.

LC-MS analysis of ChuW degradation of heme reaction mixtures showed a large amount formation of 5′-dA and SAH in the reaction (Figure 4.4).

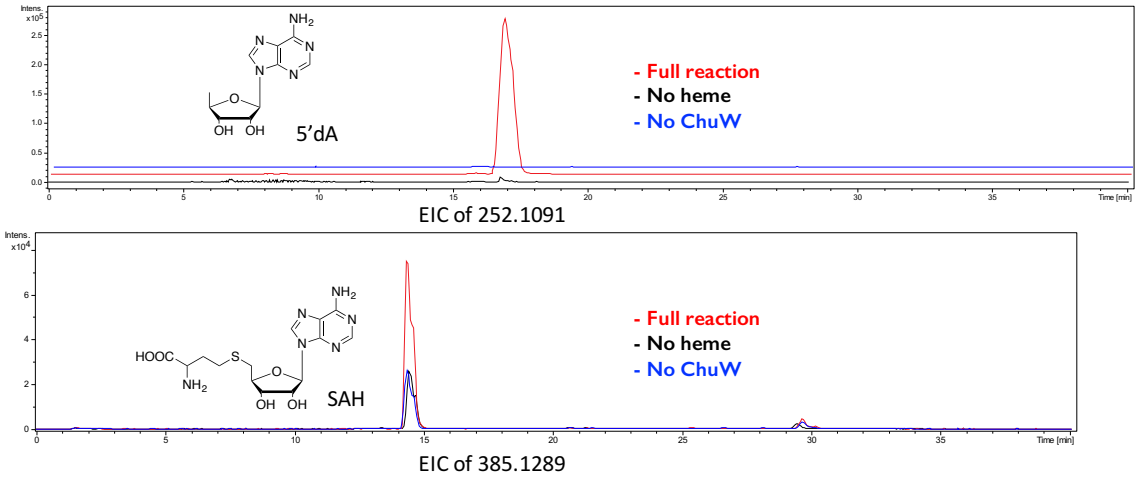


Figure 4.4 LC-MS analysis of ChuW catalyzed heme degradation. Upper panel: EICs of 5′-dA $[M+H]^+$ (252.11±0.02). Lower panel: EICs of SAH $[M+H]^+$ (385.12±0.02).

HPLC analysis of ChuW degradation of heme reaction mixtures also showed formation of a new speak under 455 nm (Figure 4.5).

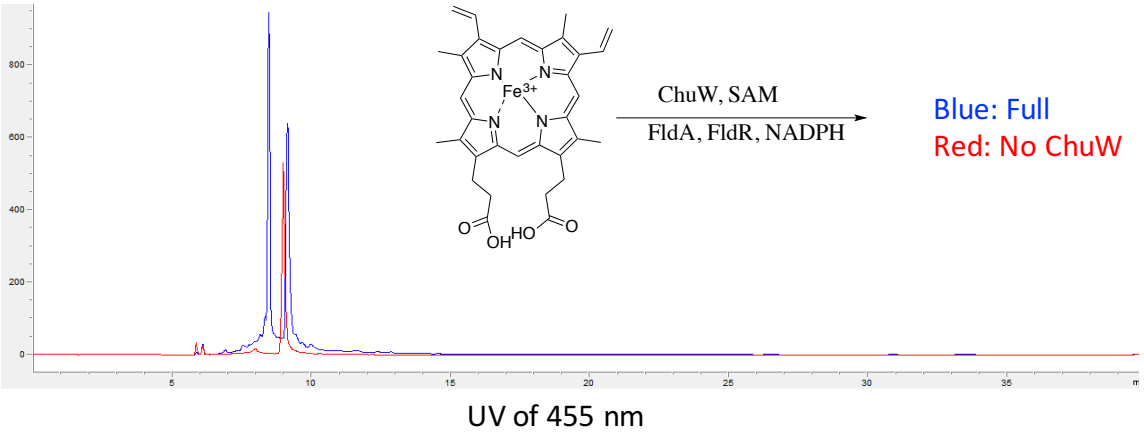


Figure 4.5 HPLC analysis of ChuW catalyzed heme degradation.

LC-MS analysis of ChuW degradation of heme reaction mixtures showed the new peak of 455 nm had a mass of 577.28, which corresponds with the mass of the ring opening product anaerobilin (Figure 4.6).

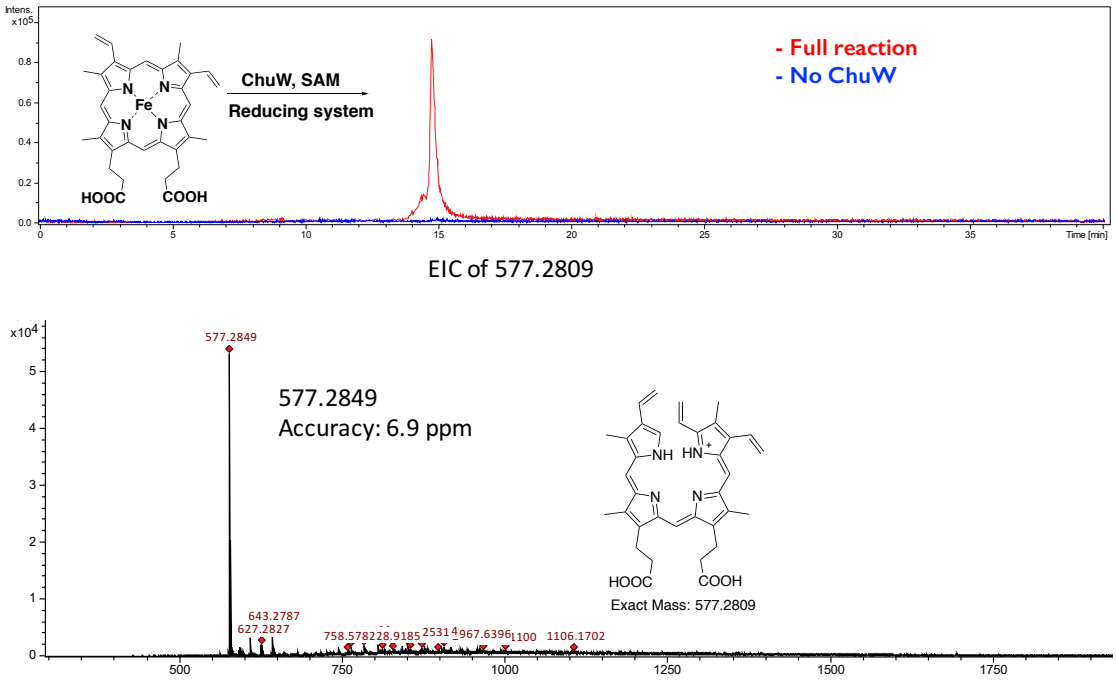


Figure 4.6 LC-MS analysis of ChuW catalyzed heme degradation. Upper panel: EICs of anaerobilin $[M+H]^+$ (577.28±0.02). Lower panel: mass spectrum of anaerobilin at 15 min.

4.2.2. Design of cyclopropyl heme analog as a radical trap for ChuW catalyzed reaction

Since there is one mass unit of 5′-dA increased from 11% to 60% when [d_3-methyl]-SAM was used in the chuW reaction, it's proposed that 5′-dA generated from the first molecular of SAM abstracts a hydrogen atom from the second molecular of

SAM. The resulting SAM radical can attack the unsaturated bond on the porphyrin ring of heme. This resulting radical species can be delocalized on the porphyrin ring to initiate the ring cleavage and the methylation. Since the ring opening of the cyclopropylcarbinyl radical is a well-recognized fast radical reaction,[79] the displacement of the vinyl group from the heme with a cyclopropyl group could be used to trap the radicals generated on the protoporphyrin ring.

The functionalization of vinyl group of tetrapyrrolic ring has been explored extensively for new photosensitizer libraries.[80] One of the useful functionalizations is a reaction with 1,3-dipolar compounds via a concerted cis-cycloaddition mechanism using diazomethane. The resulting 1-pyrazolines can be photolyzed or pyrolyzed to give cyclopropanes.[81] Cyclopropyl heme analog **4-6** is synthesized using the literature route as shown in Figure 4.7.

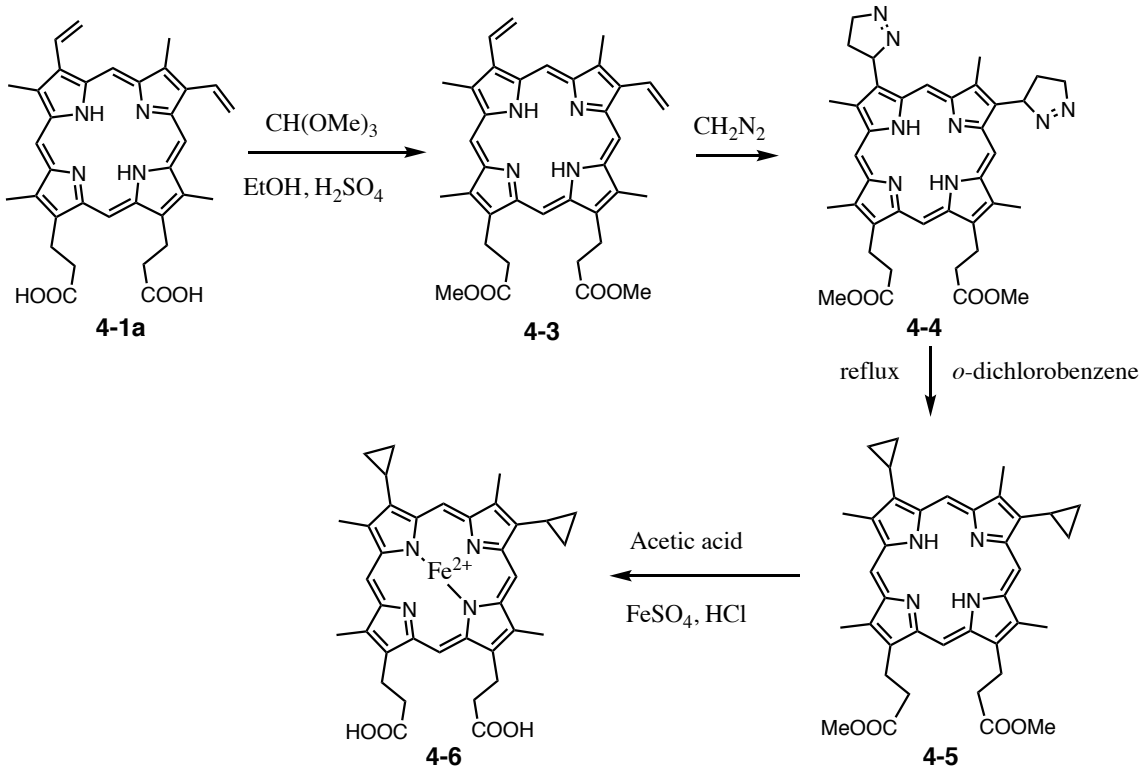


Figure 4.7 Synthetic route of cyclopropyl heme analog

4.2.3. Cyclopropyl heme analog as a substrate for ChuW

When cyclopropyl heme **4-6** was used as a substrate for ChuW with EcFldA/ EcFpr/ NADPH system as the reducing system, a new species (455nm, 780nm) was formed, which is blue shifted when compared with the anaerobilin based on the UV spectrum (Figure 4.8).

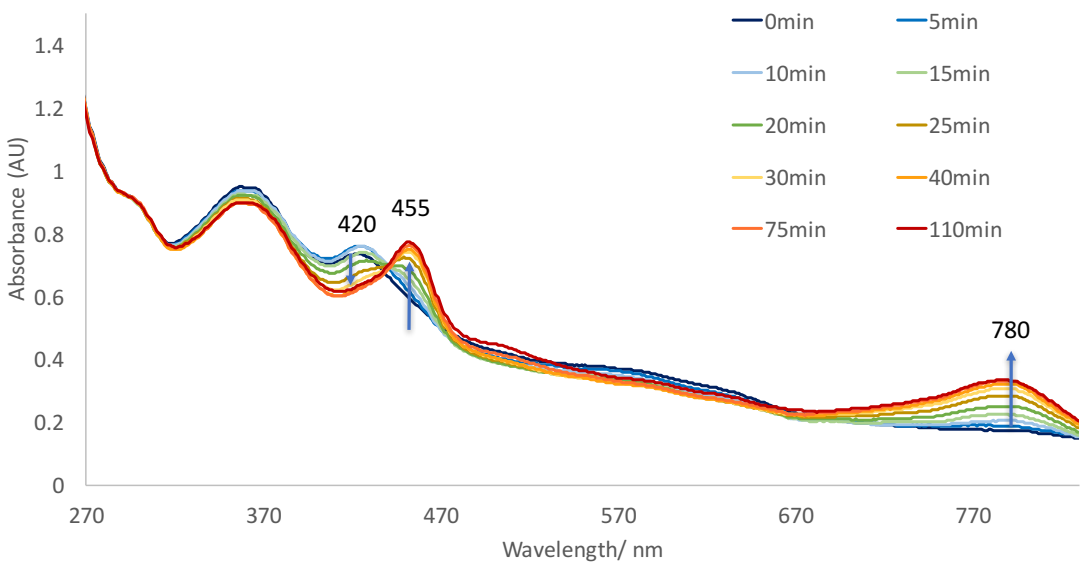


Figure 4.8 UV spectrum of the time course of the ChuW catalyzed cyclopropyl heme degradation.

LCMS analysis of ChuW degradation of cyclopropyl heme **4-6** reaction mixtures showed a large amount formation of 5′-dA in the reaction, which suggests that ChuW takes cyclopropyl heme as a substrate (Figure 4.9).

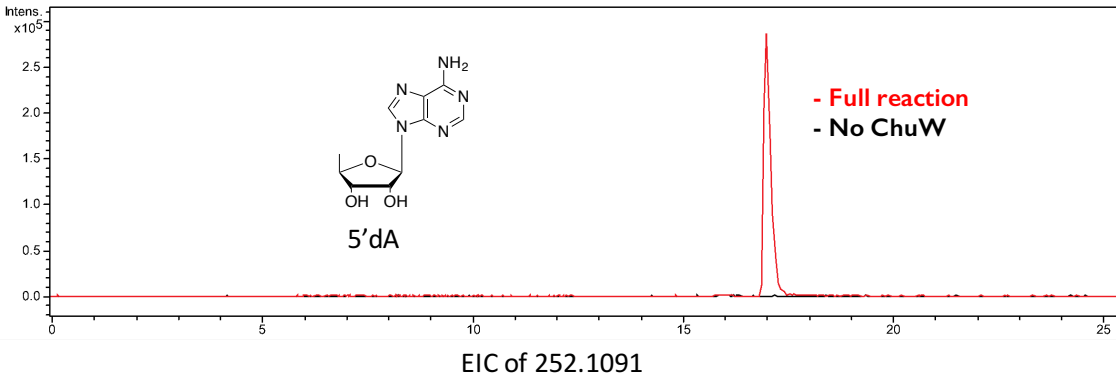


Figure 4.9 LC-MS analysis of ChuW catalyzed cyclopropyl heme **4-6** degradation. EICs of 5′-dA $[M+H]^+$ (252.11±0.02).

HPLC analysis of ChuW degradation of cyclopropyl heme **4-6** reaction mixtures also showed the formation of new species which have an absorbance around 500nm. The high wavelength products are not stable under air and light condition. The 500nm products decrease over time (Figure 4.10).

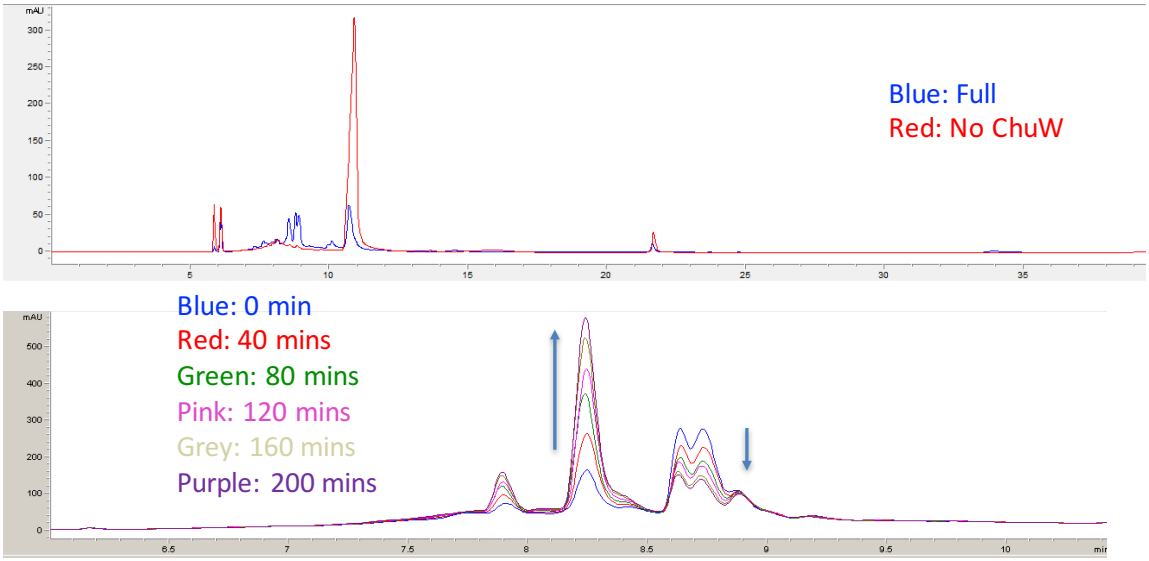


Figure 4.10 HPLC analysis of ChuW catalyzed cyclopropyl heme degradation under UV of 500 nm.

LCMS analysis of new compounds with high wavelength showed a new peak of 500 nm had a mass of 605.31, which corresponds with the mass of the methylated ring opening product (Figure 4.11).

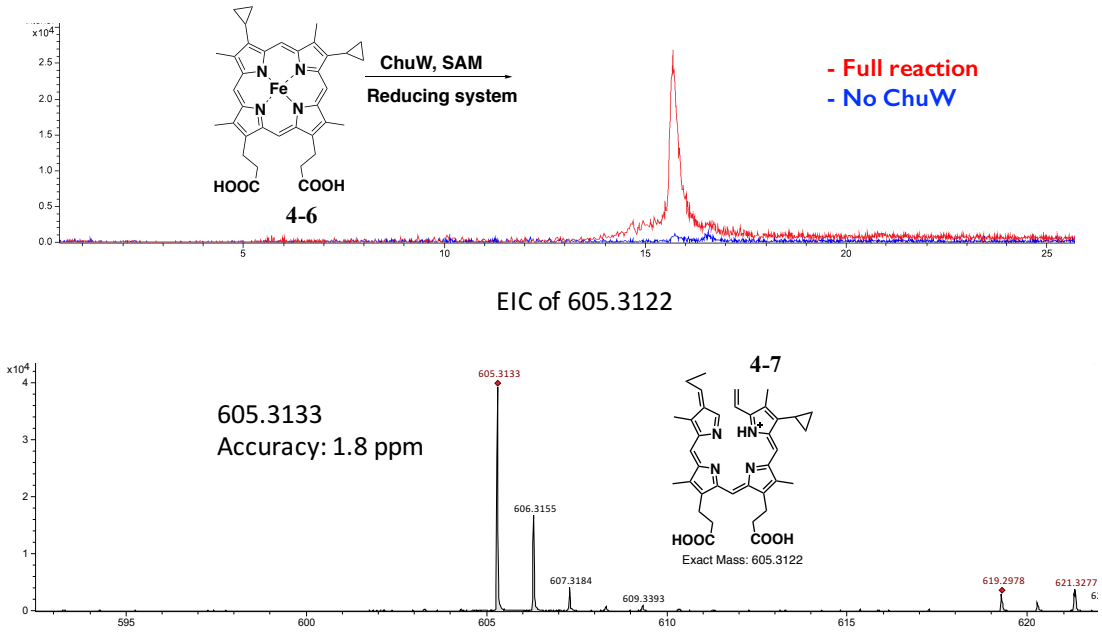


Figure 4.11 LC-MS analysis of ChuW catalyzed cyclopropyl heme degradation. Upper panel: EICs of $[M+H]^+$ (605.31±0.02). Lower panel: mass spectrum of degradation product at 15 min.

When CD_3-SAM or $^{13}CH_3$-SAM was used in the ChuW catalyzed cyclopropyl heme degradation reaction, there were two mass units and one mass unit increase in the 605.31 peak respectively (Figure 4.12). This suggests that the methyl group is transferred to the new product from SAM. However, this compound is also not stable under the light condition. The high wavelength peak will disappear when it's exposed to light.

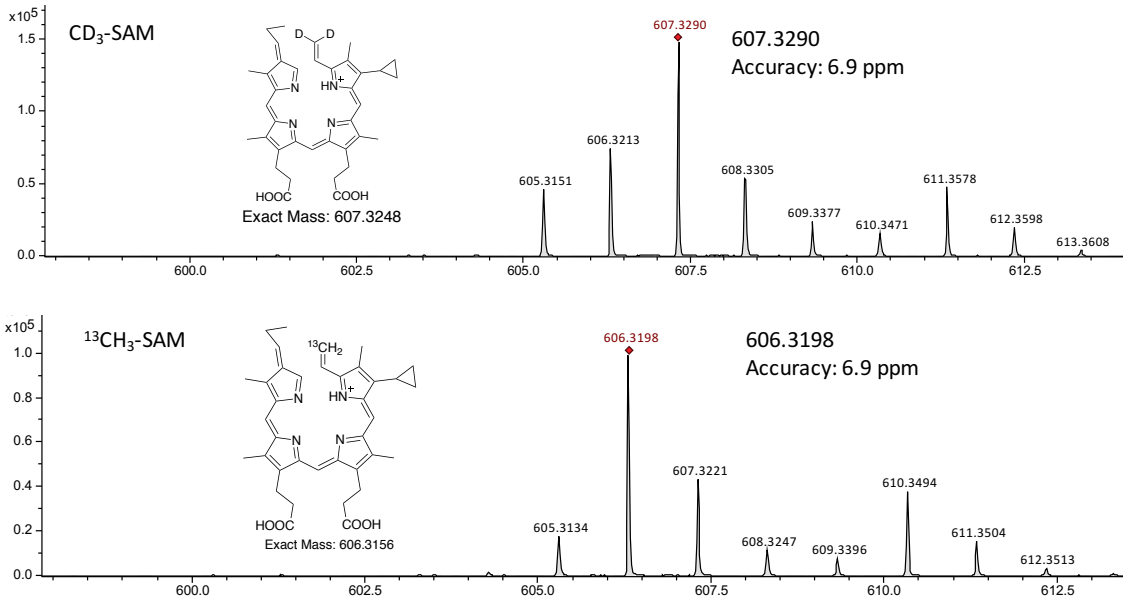


Figure 4.12 Mass spectrum of ChuW catalyzed cyclopropyl heme degradation product at 15 min with (A) CD_3-SAM or (B) $^{13}CH_3$-SAM.

4.2.4. Detection of protoporphyrin ring opening product in ChuW catalyzed cyclopropyl heme 4-6 degradation

Under the UV wavelength of 390 nm, there was a new peak formation around 34 min in HPLC chromatography, which has a similar UV spectrum as the protoporphyrin (Figure 4.13). Because of the highly conjugated π-electron systems of the tetrapyrrole ring, the porphyrins have characteristic UV-visible spectra that consist of two distinct region regions: in the near ultraviolet and the visible region.[82] The porphyrins display extreme intense bands, the so-called Soret or B-bands in the 380–500 nm range with molar extinction coefficients of 10^5 M^{-1} cm^{-1}. Moreover, at longer wavelengths, in the 500–750 nm range, the porphyrins spectra contain a set of weaker, but still considerably intense Q bands with molar extinction coefficients of 10^4 M^{-1} cm^{-1}. As for the new peak

generated in the ChuW catalyzed cyclopropyl heme **4-6** degradation, it has a large absorbance around 390 nm in the Soret bands region, and some weaker absorbance in the 500, 530, 570 nm in the Q-bands region (Figure 4.13). This suggests that this new peak should still maintain the intact porphyrin ring. This product is considered stable under the light condition for several hours (Figure 4.14).

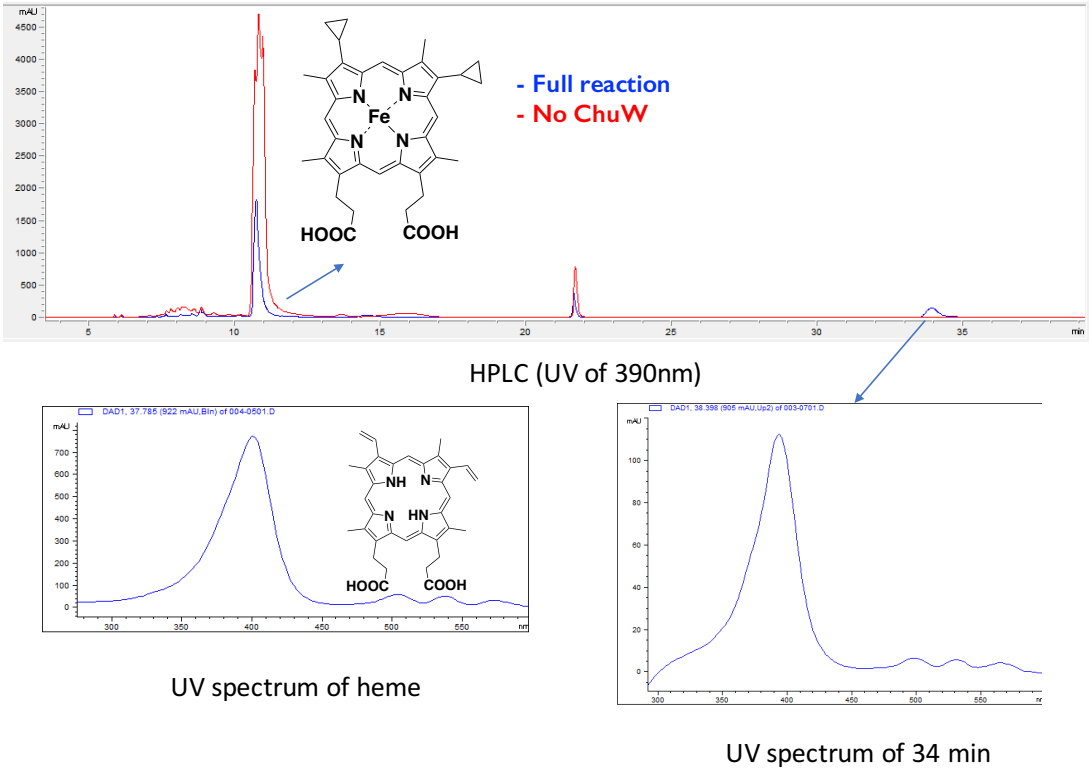


Figure 4.13 HPLC analysis of ChuW catalyzed cyclopropyl heme degradation. Upper panel: HPLC UV trace of 390 nm. Lower panel left: UV spectrum of heme; Lower panel right: UV spectrum of product at 34 min.

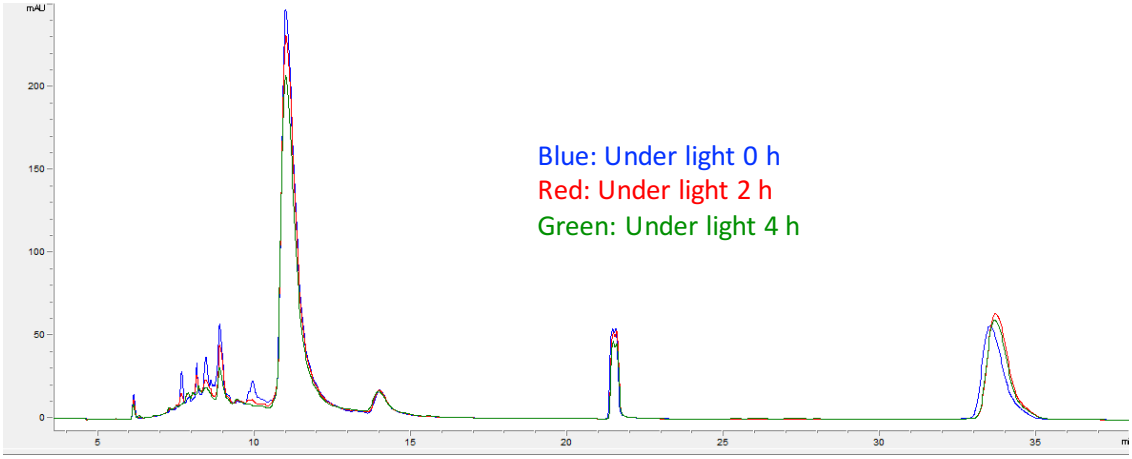


Figure 4.14 HPLC analysis of stability of products in ChuW catalyzed cyclopropyl heme degradation over time under 400 nm

LC-MS analysis of the new compound showed it had a mass of 591.3011, which corresponds with the mass of cyclopropyl heme analog without the iron (Figure 4.15B). When CD_3-SAM was used in the ChuW catalyzed reaction, 60% of the product have one mass unit increased (Figure 4.15C). To further explore whether the cyclopropyl side chain is open, the porphyrin product was purified from HPLC and was hydrogenated in a short amount of time. A new peak with a mass of 593.31 was formed (Figure 4.15D). This suggests that there is a double bound on the side chain of the porphyrin ring and probably is the cyclopropyl ring opening product.

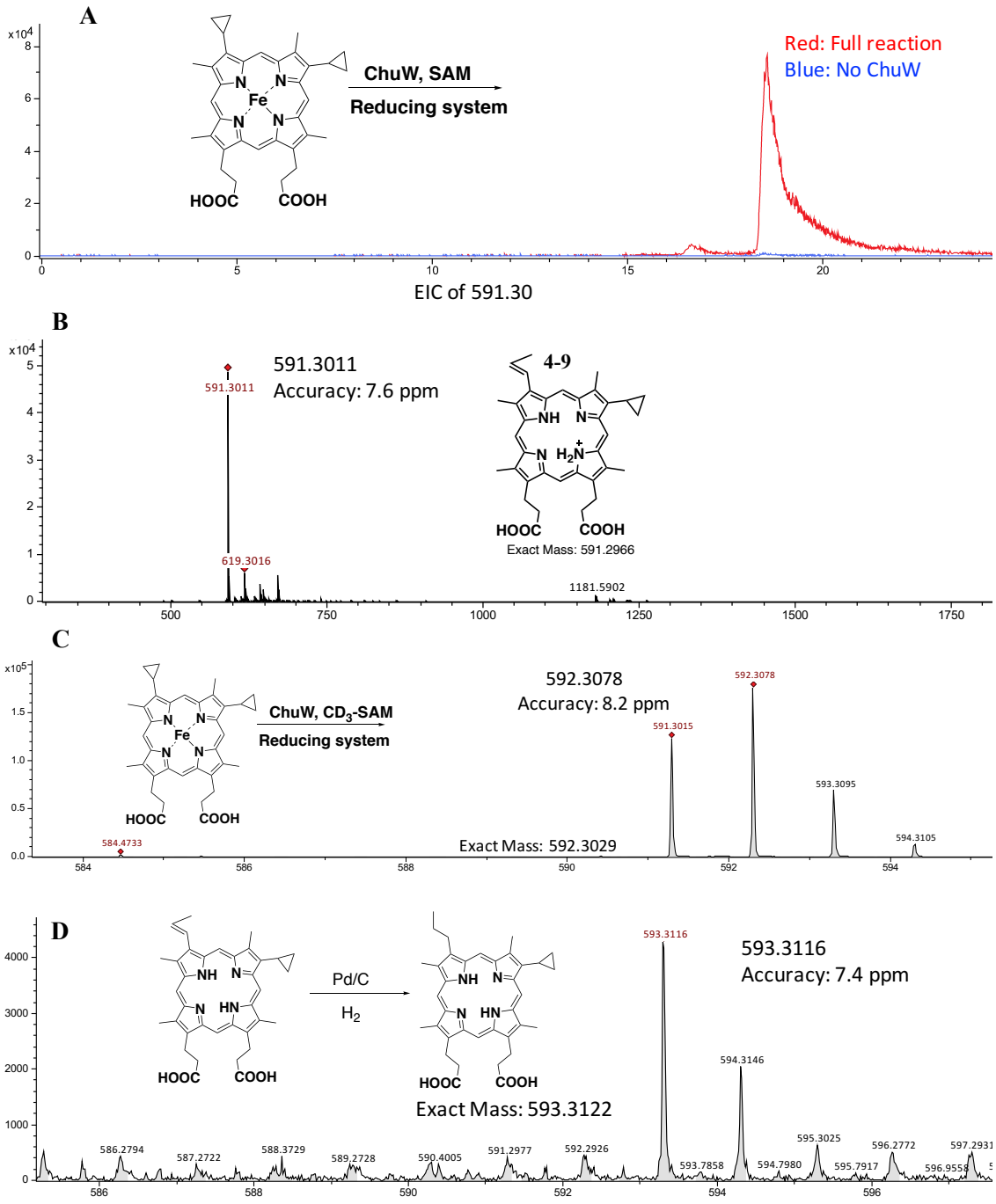


Figure 4.15 LC-MS analysis of ChuW catalyzed cyclopropyl heme degradation. (A) EICs of $[M+H]^+$ (591.30±0.02). (B) Mass spectrum of the porphyrin degradation product at 19 min. (C) Mass spectrum of the porphyrin degradation product when the reaction used CD_3-SAM as cofactor. (D) Mass spectrum of the porphyrin degradation product after hydrogenation reaction.

4.2.5. Detection of SAM adducts in ChuW catalyzed cyclopropyl heme degradation

LC-MS analysis of the unstable high wavelength products showed a compound had a mass corresponding to the crosslinking species **4-10** between cyclopropyl heme derived product and SAM, which were absent in the control (Figure 4.16). Another compound also had a mass corresponding to the crosslinking species **4-11** between cyclopropyl heme derived product and methylthioadenosine (Figure 4.17). It is known that SAM can degrade to L-homoserine lactone (HSL) and 5-methylthioadenosine (MTA) under high and low pH condition.[83] After purification and acid treatment of the reaction mixture, SAM crosslinking species **4-10** can be degraded to MTA crosslinking species **4-11**.

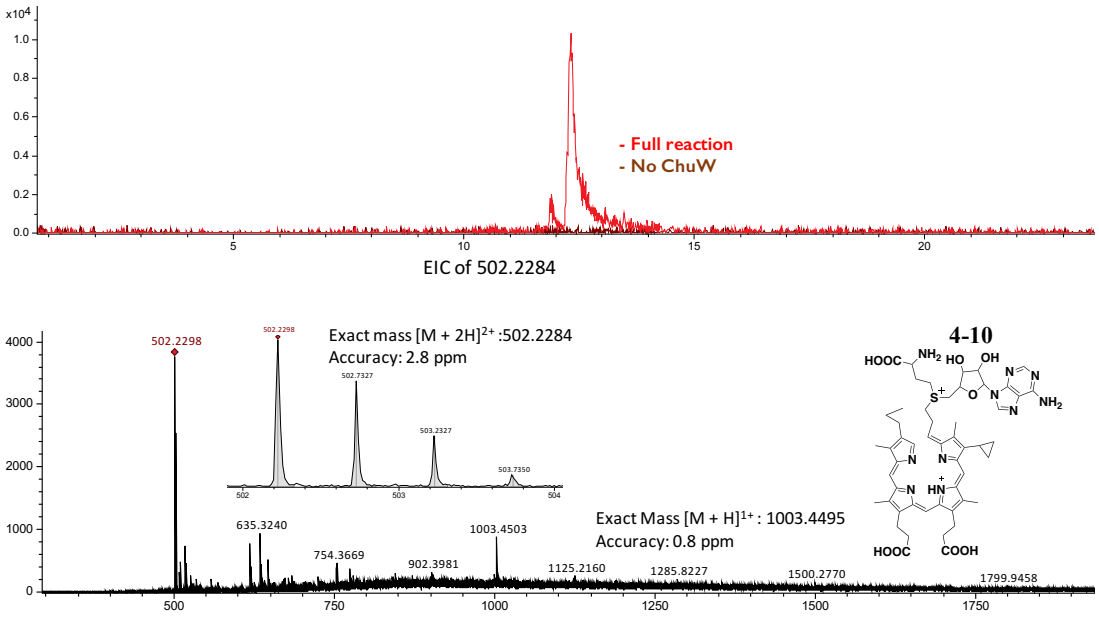


Figure 4.16 LC-MS analysis of ChuW catalyzed cyclopropyl heme degradation. Upper panel: EICs of $[M+2H]^{2+}$ (502.22±0.02). Lower panel: mass spectrum of the porphyrin cross-linking species **4-10** at 12.5 min.

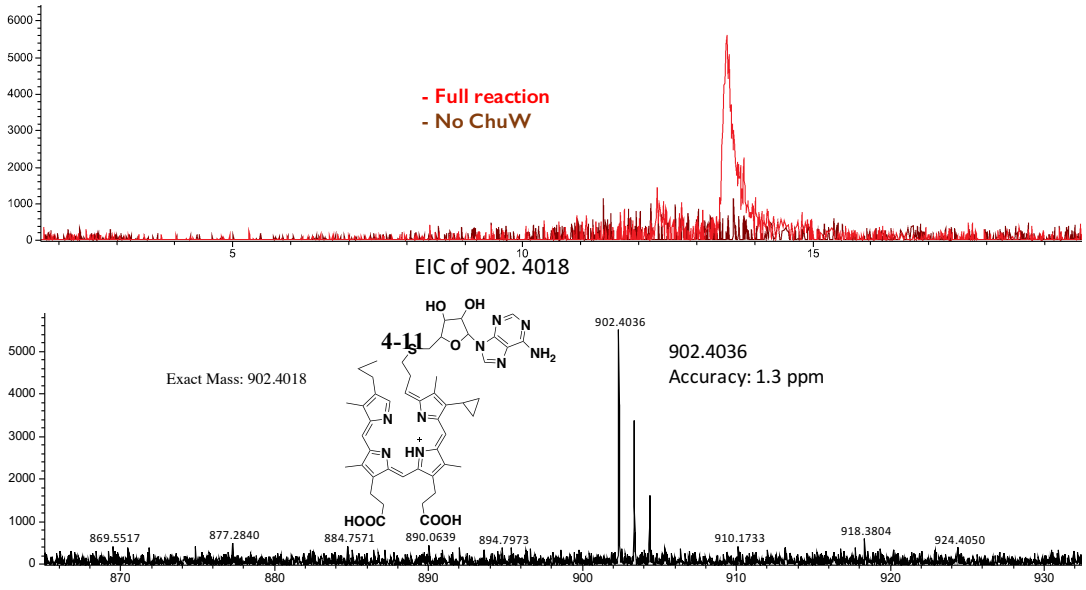


Figure 4.17 LC-MS analysis of ChuW catalyzed cyclopropyl heme degradation. Upper panel: EICs of $[M+H]^+$ (902.40±0.02). Lower panel: mass spectrum of the degradation porphyrin cross-linking species **4-11** at 13.5 min.

When CD_3-SAM was used in the reaction, the majority of the crosslinking species have 2-4 mass unit increased for the +1-charged compound (Figure 4.18, 4.19). Similarly, when $^{13}CH_3$-SAM was used in the reaction, the majority of the crosslinking species have 2 mass unit increased for SAM adduct **4-10** (Figure 4.18, 4.19). MS/MS fragmentations of the crosslinking species also match with two methylene groups from SAM get incorporated into the product (Figure 4.20).

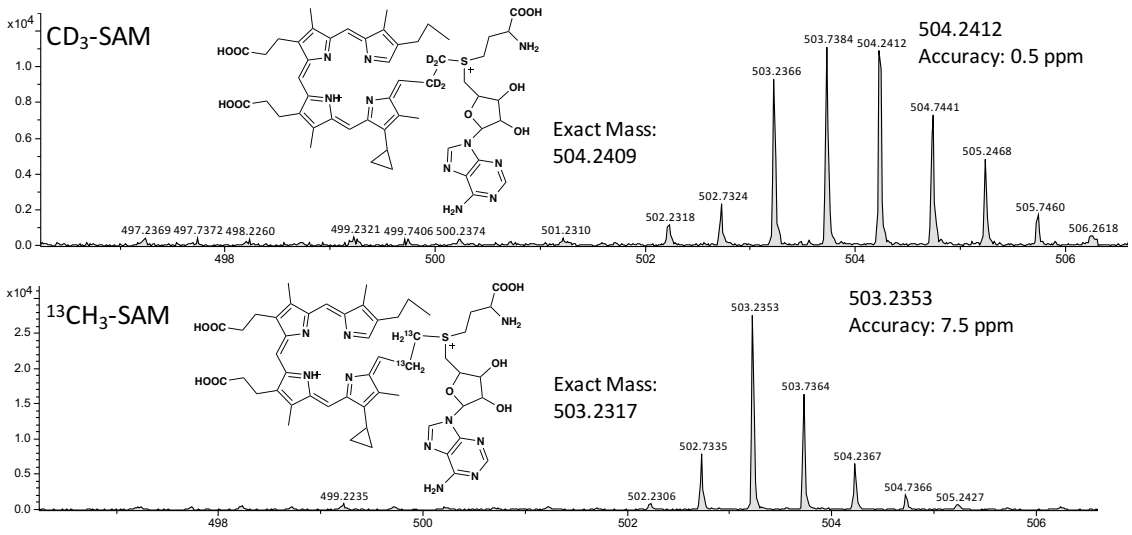


Figure 4.18 Mass spectrum of the porphyrin cross-linking species (503.22) when ChuW used (upper panel) CD_3-SAM, (lower panel) $^{13}CH_3$-SAM as cofactor.

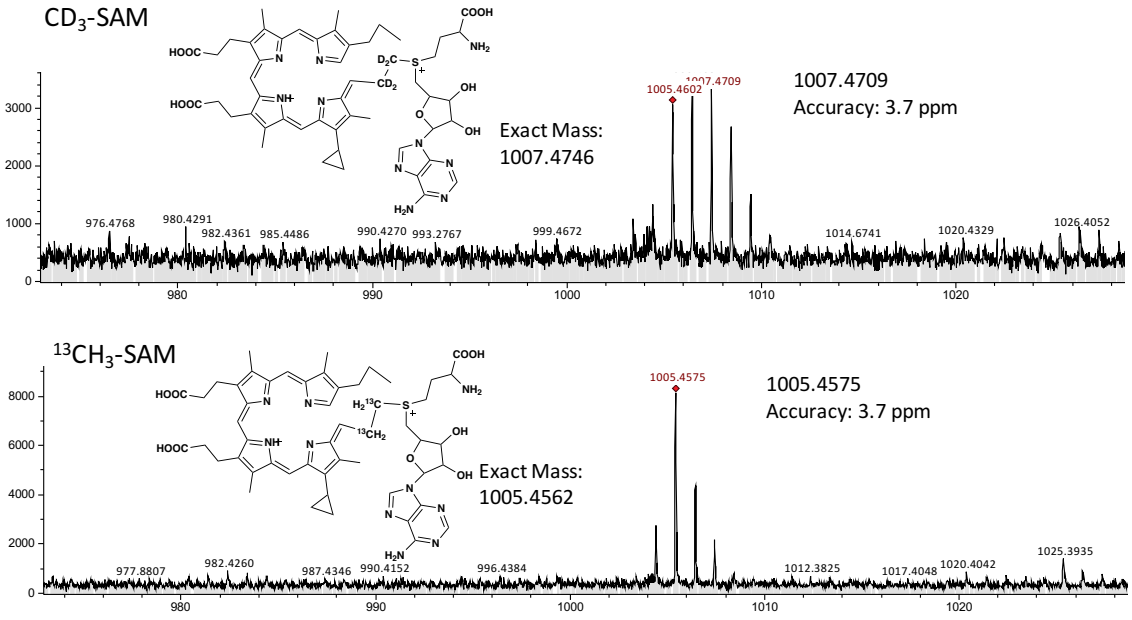


Figure 4.19 Mass spectrum of the porphyrin cross-linking species **4-10** (1003.45) when ChuW used (upper panel) CD_3-SAM, (lower panel) $^{13}CH_3$-SAM as the cofactor.

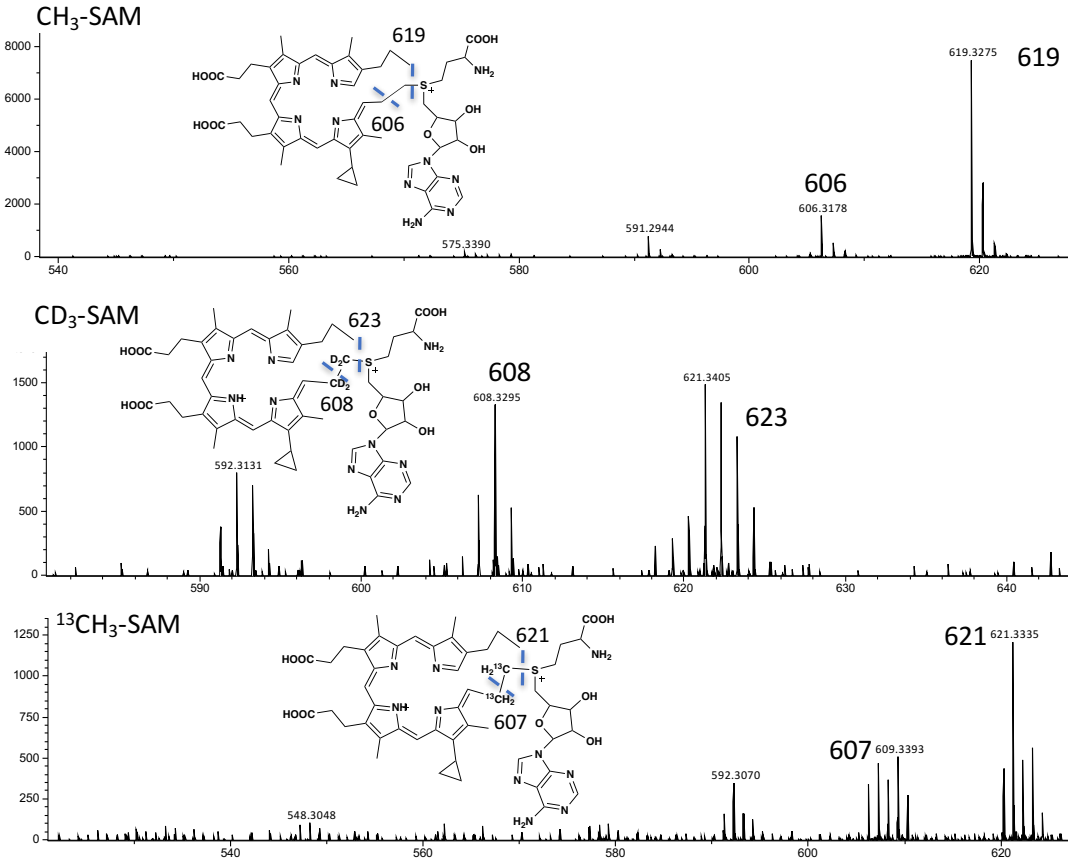


Figure 4.20 Mass fragmentation of the porphyrin cross-linking species (502.23) when ChuW used (upper panel) CH_3-SAM, (middle panel) CD_3-SAM, (lower panel) $^{13}CH_3$-SAM as cofactor.

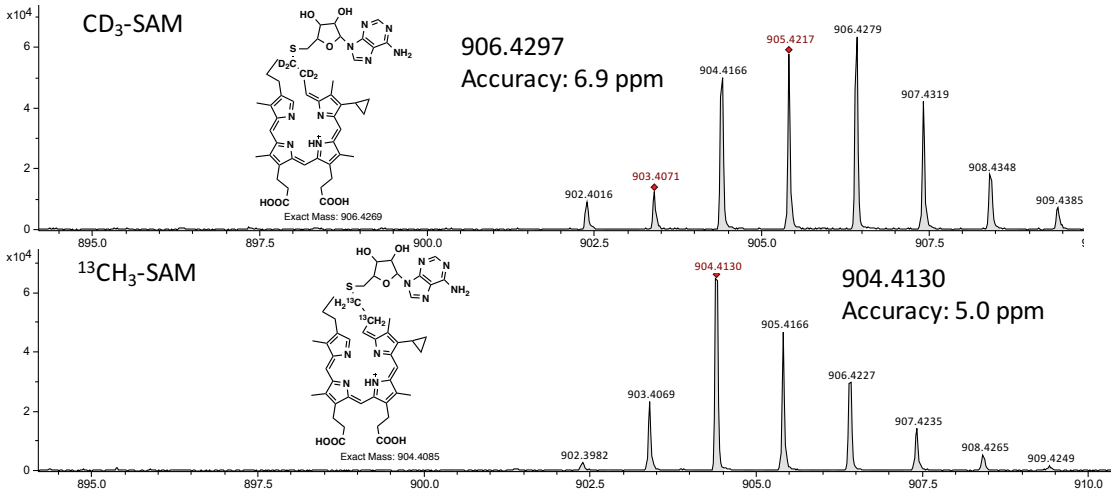


Figure 4.21 Mass spectrum of the degradation porphyrin cross-linking species **4-11** (902.40) when the ChuW used (upper panel) CD_3-SAM, (lower panel) $^{13}CH_3$-SAM as cofactor.

When CD_3-SAM was used as the cofactor and titanium citrate was used as the reducing agent in the ChuW catalyzed reaction, the [A+1] peak for 5'-dA increased from 13% to 98% of [A], the [A+2] peak for 5'-dA increased from 1.8% to 31% of [A] (Figure 4.23). It's been shown in the literature that 5'-dA radical can abstract a hydrogen atom from the second molecule of SAM (Figure 4.22). Thus, it's expected that there will be 1 mass unit increased for the 5'-dA when CD_3-SAM was used. However, the increase of [A+2] peak indicated that the ring opening radical probably could abstract a hydrogen atom from 5'-dA, which then can abstract a hydrogen atom again from CD_3-SAM.

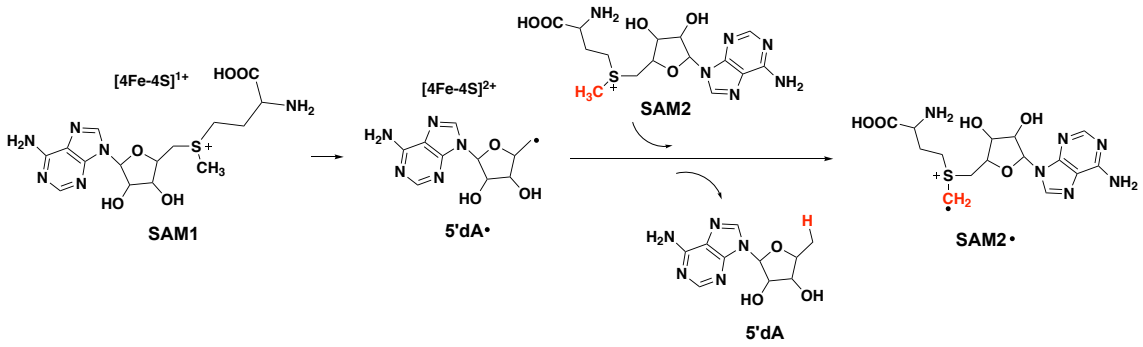


Figure 4.22 The general scheme of SAM radical generated in the Class C radical SAM enzyme.

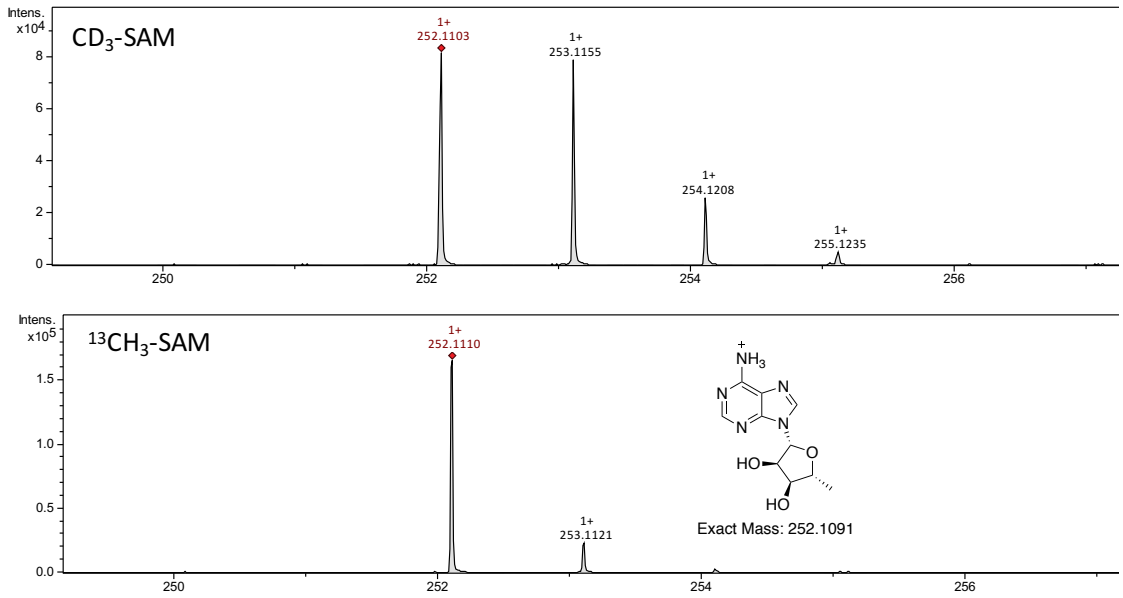


Figure 4.23 Mass spectrum of 5'-dA when ChuW used (upper panel) CD_3-SAM, (lower panel) $^{13}CH_3$-SAM as the cofactor.

4.2.6. ChuW catalyzed cyclopropyl protoporphyrin degradation

Since protoporphyrin can also be a substrate for ChuW, cyclopropyl protoporphyrin **4-12** was synthesized and was tested whether it could also be a substrate

for ChuW. LC-MS analysis of ChuW degradation of cyclopropyl protoporphyrin reaction mixtures showed a large amount formation of 5′-dA in the reaction, which suggests that ChuW takes cyclopropyl protoporphyrin as a substrate (Figure 4.24).

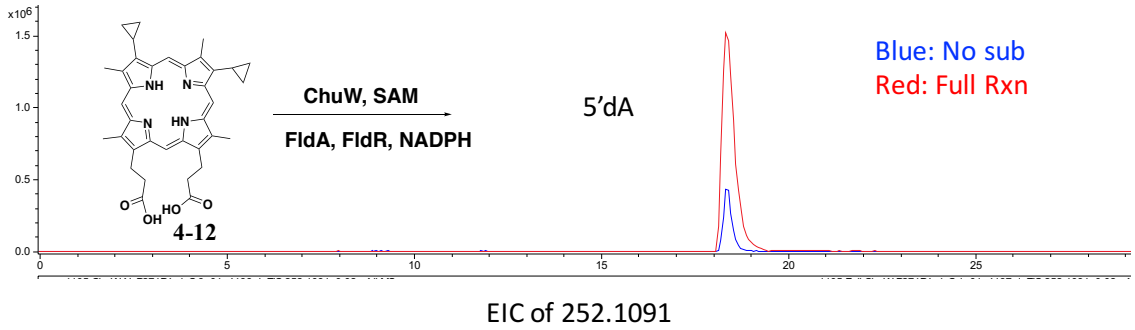
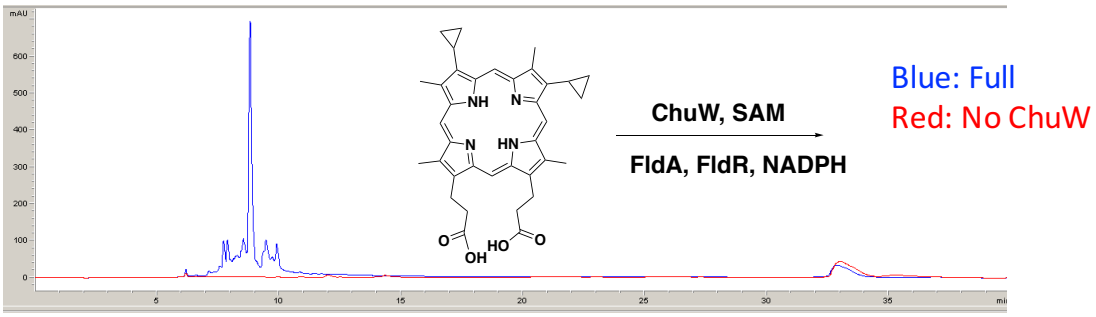


Figure 4.24 LC-MS analysis of ChuW catalyzed cyclopropyl protoporphyrin degradation. EICs of 5'-dA $[M+H]^+$ (252.10±0.02).

HPLC analysis of ChuW degradation of cyclopropyl protoporphyrin reaction mixtures also showed the formation of new species which have absorbance around 455 nm (Figure 4.25). The high wavelength products are not stable under air and light condition. The 500nm products will decrease over time, which is similar to the case when cyclopropyl heme **4-6** is used as a substrate for ChuW.



UV of 455 nm

Figure 4.25 HPLC analysis of ChuW catalyzed cyclopropyl protoporphyrin **4-12** degradation under 455 nm UV.

Similarly, as the cyclopropyl heme substrate, LC-MS analysis of high wavelength compounds showed a mass of 605.31, which corresponds with the mass of the ring opening product **4-7** (Figure 4.26). What's more, a new peak with a mass of 607.32 was also detected, which corresponds with the further reduced form of the ring opening product (Figure 4.27).

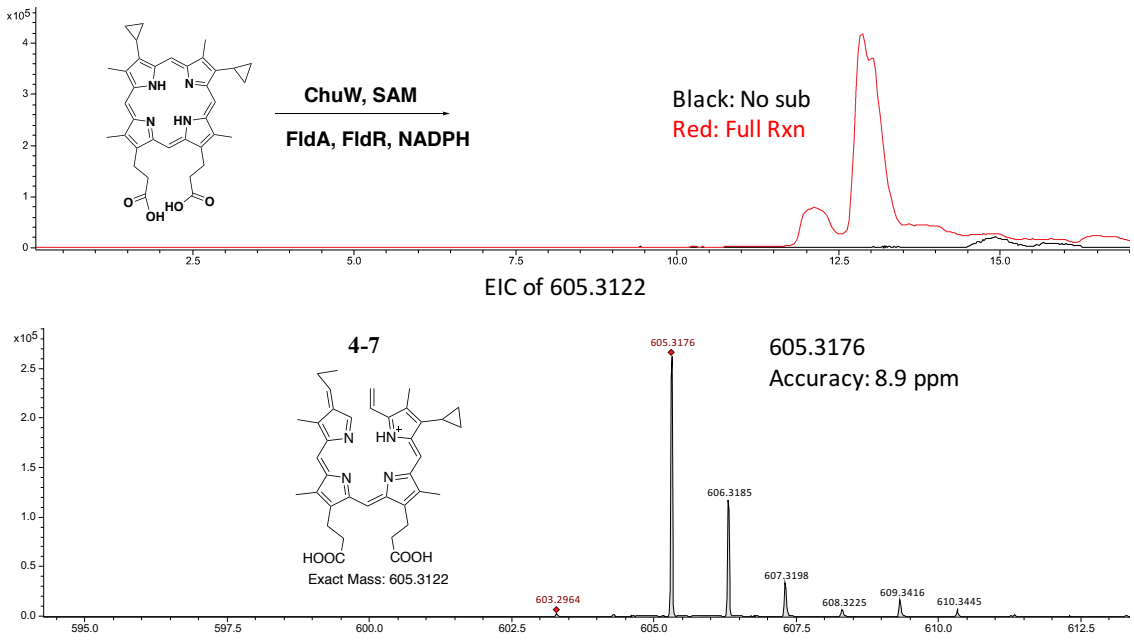


Figure 4.26 LC-MS analysis of ChuW catalyzed cyclopropyl protoporphyrin degradation. Upper panel: EICs of the ring cleavage product $[M+H]^+$ (605.31±0.02). Lower panel: mass spectrum of the ring cleavage product **4-7** at 13 min.

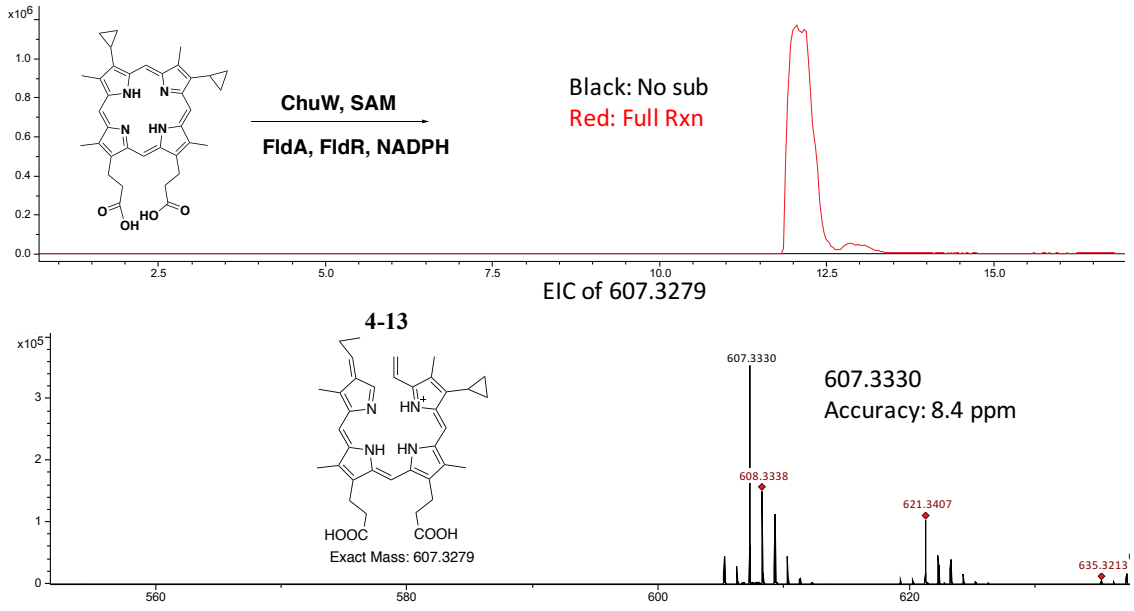


Figure 4.27 LC-MS analysis of ChuW catalyzed cyclopropyl protoporphyrin degradation. Upper panel: EICs of the reduced form of ring cleavage product $[M+H]^+$

(607.33±0.02). Lower panel: mass spectrum of the reduced form of ring cleavage product **4-13** at 12 min.

Mass corresponding with the crosslinking species between cyclopropyl ring product and SAM was also detected (Figure 4.28). When CD_3-SAM was used in the reaction, the majority of the crosslinking species 2-4 mass units increased for the SAM adduct 4-10 (Figure 4.29). When D_4-SAM was used in the reaction, the crosslinking species also have 4 mass units increased for the SAM adduct 4-10 (Figure 4.29). The degradation and reduced form of the crosslinking species was also detected (Figure 4.30).

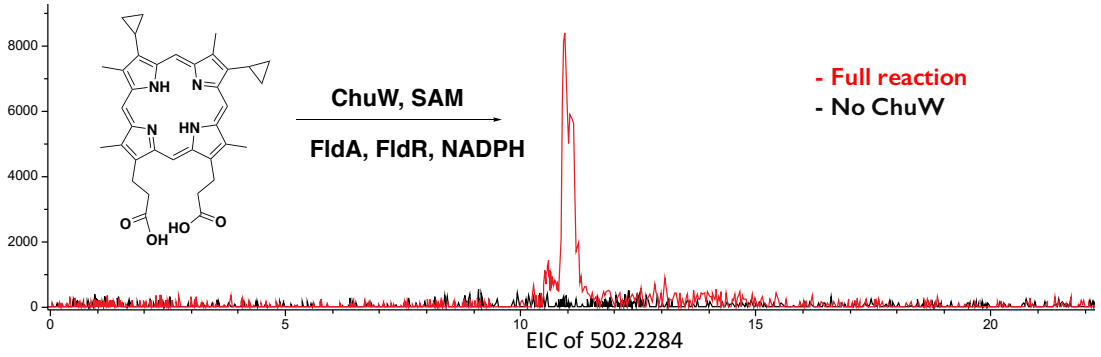


Figure 4.28 LC-MS analysis of ChuW catalyzed cyclopropyl protoporphyrin degradation. EICs of the porphyrin cross-linking species $[M+H]^+$ (502.23±0.02).

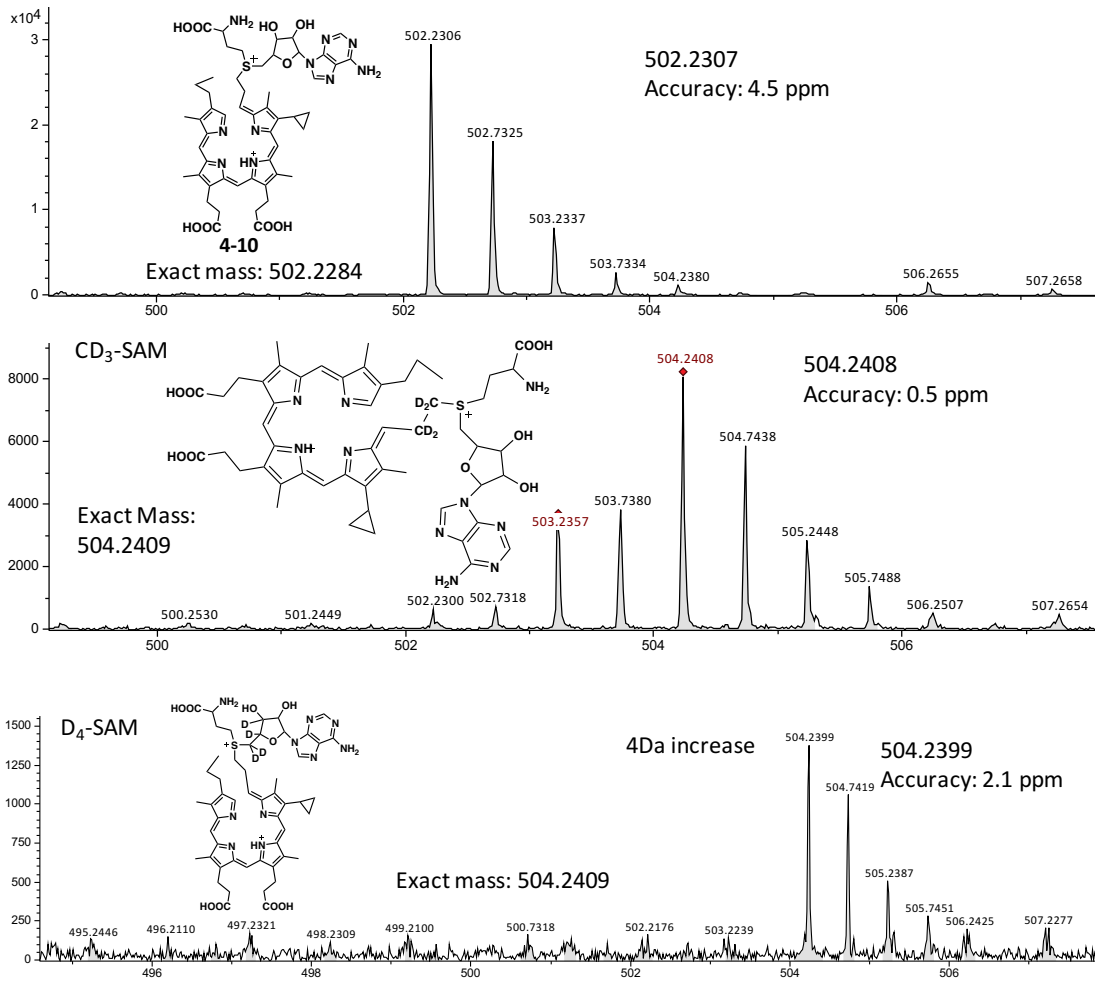


Figure 4.29 Mass spectrum of the porphyrin cross-linking species (502.23) when the ChuW used (upper panel) CH_3-SAM, (middle panel) CD_3-SAM, (lower panel) D_4-SAM as cofactor.

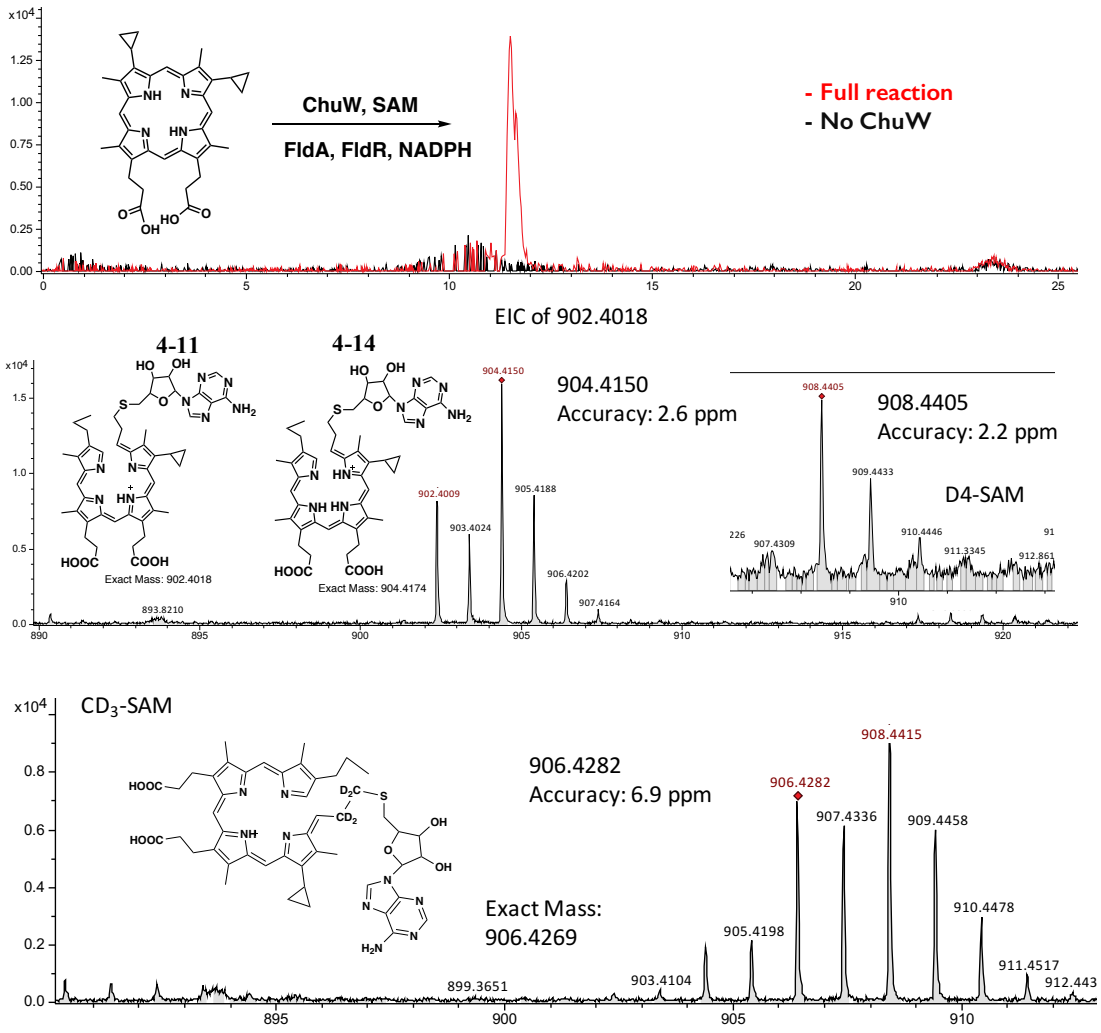


Figure 4.30 LC-MS analysis of ChuW catalyzed cyclopropyl protoporphyrin degradation. Upper panel: EICs of the degradation product of porphyrin cross-linking species $[M+H]^+$ (902.40±0.02). Middle panel: mass spectrum of the degradation product of porphyrin cross-linking species at 11.5 min; Zoon in graph: mass spectrum of the degradation product of porphyrin cross-linking species at 11.5 min when D_4-SAM was used as cofactor; Lower panel: mass spectrum of the degradation product of porphyrin cross-linking species at 11.5 min when CD_3-SAM was used as cofactor.

When a strong reducing reagent, such as titanium citrate, was used in the ChuW catalyzed reaction, a crosslinking species **4-15** was detected with just one methylene group from SAM got incorporated (Figure 4.31). Based on the UV spectrum, the

crosslinking species most likely contains the ring opening form of the porphyrin ring. When CD_3-SAM was used in the reaction, the majority of the crosslinking species have 2 mass units increased, which suggested that there was one methylene group from the methyl group of SAM (Figure 4.32).

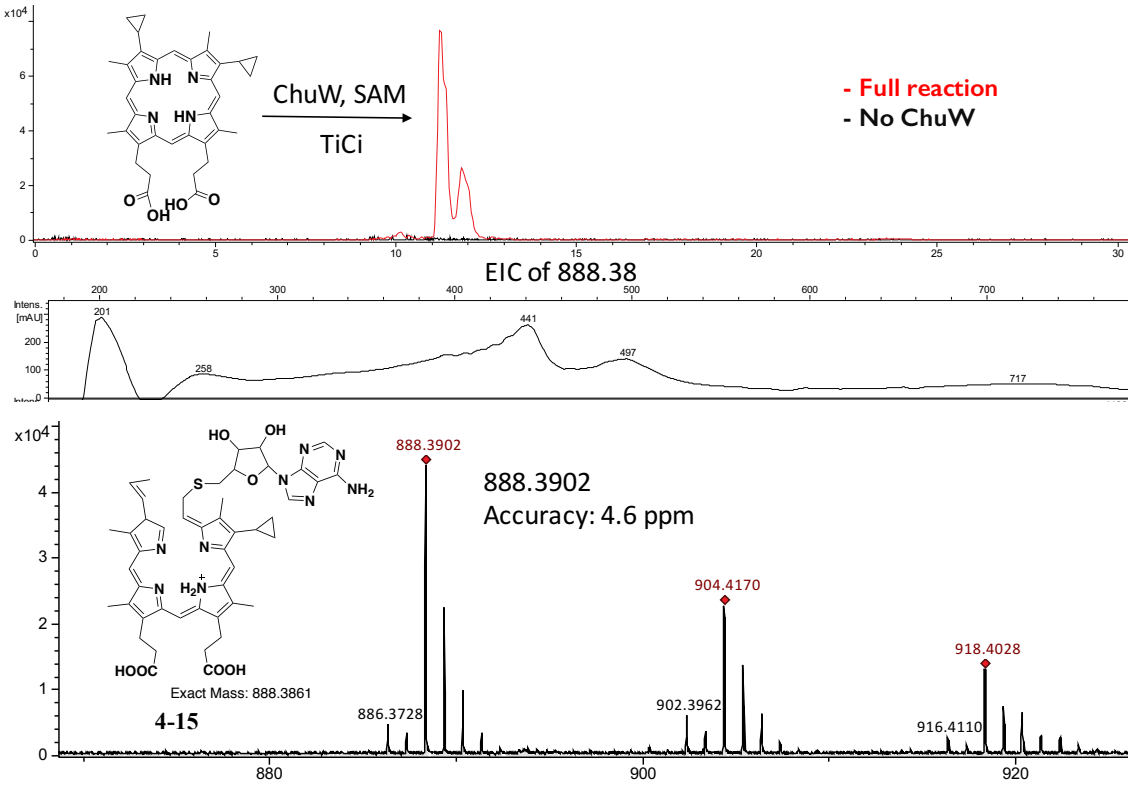


Figure 4.31 LC-MS analysis of ChuW catalyzed cyclopropyl protoporphyrin degradation using titanium citrate as reducing agent. Upper panel: EICs of the cross-linking species $[M+H]^+$ (888.38±0.02). Middle panel: UV spectrum of the cross-linking species at 11 min; Lower panel: mass spectrum of the cross-linking species at 11 min.

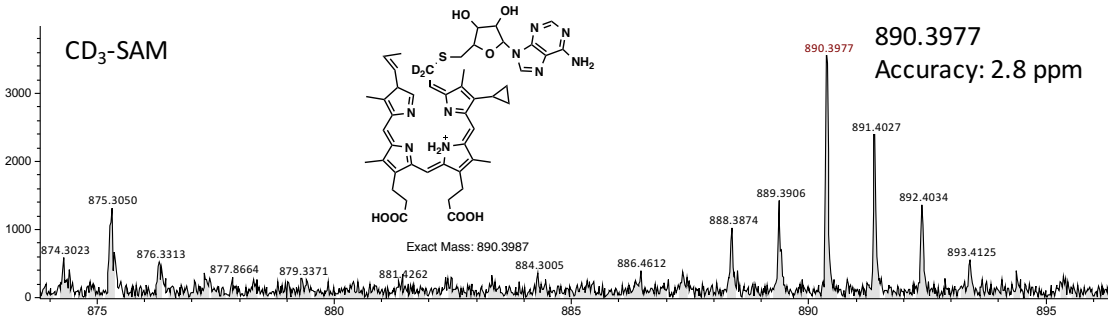


Figure 4.32 Mass spectrum of the cross-linking species in ChuW catalyzed cyclopropyl protoporphyrin degradation using titanium citrate as reducing agent.

4.2.7. Proposed mechanisms for the ring opening product of ChuW catalyzed cyclopropyl substrates reaction

Based on the result that 5′-dA radical abstracts a hydrogen atom from a methyl group of SAM and the methylene group from SAM is transferred to the ring opening tetrapyrrole product in the ChuW catalyzed reaction, several mechanisms can be proposed for the formation of protoporphyrin ring opening product **4-9**. Here, two mechanisms are shown in Figure 4.33 and Figure 4.34. After 5′-dA radical abstracts a hydrogen atom from a methyl group of SAM, the SAM radical attacks the unsaturated bond of the **4-6** to form **4-21**. In pathway a, deprotonation and tautomerization of **4-21** lead SAM left as a ylide, and the resulting radical delocalize on the porphyrin and induces the ring opening of the cyclopropyl group on porphyrin A ring to give **4-23**, which can be quenched by abstracting a hydrogen atom from SAM (Figure 4.33). This can account for the deuterium incorporation of the porphyrin product when CD_3-SAM was used. For pathway b, the early step is the same as the pathway a, which forms the intermediate **4-21**. The resulting radical induces the ring opening of the cyclopropyl

group on porphyrin B ring. The radical on **4-21** attacks back on the unsaturated bond of pyrrole to form a five-member ring on porphyrin B ring. Then **4-28b** goes through a beta-scission to give the **4-29**. Both **4-9** and **4-29** have the same mass and can be reduced under hydrogenation. NMR characterization of the enzymatic product will help to distinguish which ring is open. Since porphyrin with 18 π electron core is considered more stable than anaerobilin under light condition, it's possible to purify compound using solid phase extraction followed by HPLC in a short amount of time. However, since the yield of enzymatic product is low and the purification steps under dark are tedious, there is not enough compound for NMR characterization yet. More compounds have to be purified in the future in order to obtain an NMR spectrum. Reference compounds **4-9** and **4-29** may also be made from irradiation of cyclopropyl protoporphyrin **4-12**.

As for the formation of SAM crosslinking species **4-10**, **4-11**, the early steps are the same as the formation of porphyrin product. However, after the SAM radical attacks the unsaturated bond of the porphyrin ring, instead of SAM leaving as ylide, the radical goes through β - scission to cleave the porphyrin ring. The pyrrole racial **4-26** gets trapped by the ring opening of side chain cyclopropyl group. After tautomerization, SAM leaves as SAH. The resulting ring opening radical **4-27** can abstract a hydrogen atom from 5′-dA, which can again abstract a hydrogen atom from SAM. Then the SAM radical can attack the methylene group of the ring opening product **4-7**, which further gets quenched by obtaining an electron and a proton to give the crosslinking species **4-10**. Under current purification condition, **4-10** can be degraded to **4-11**. When

cyclopropyl protoporphyrin **4-12** was used as a substrate, and titanium citrate was used as a reducing agent, the intermediate SAM adduct **4-15** can be detected from the reduction of **4-26**.

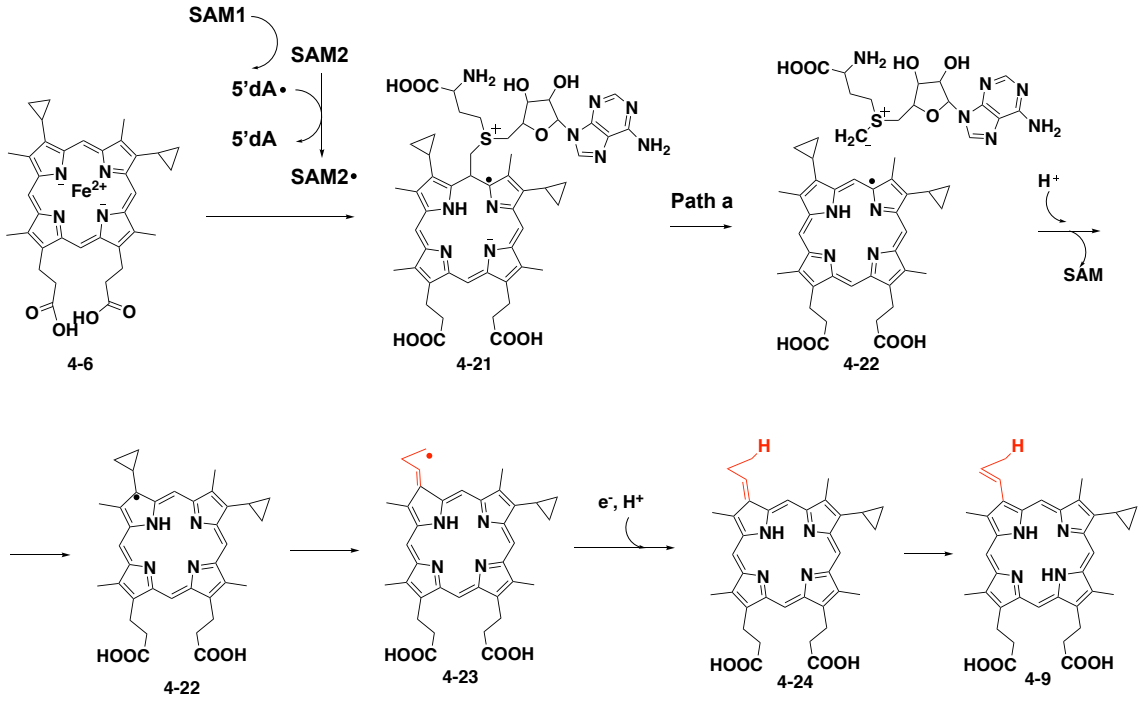


Figure 4.33 Proposed mechanism 1 for how cyclopropyl heme substrate analog **4-6** trap radical intermediate by formation of porphyrin ring opening product in the ChuW catalyzed reaction.

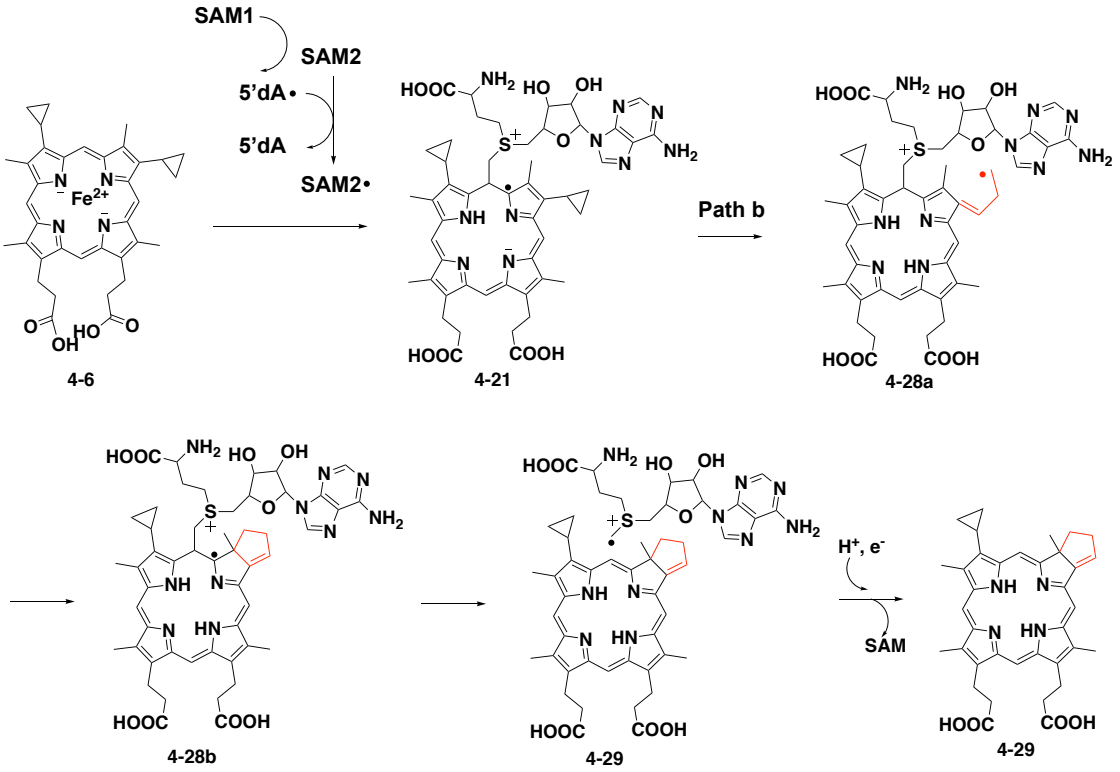


Figure 4.34 Proposed mechanism 2 for how cyclopropyl heme substrate analog **4-6** trap radical intermediate by formation of porphyrin ring opening product in the ChuW catalyzed reaction.

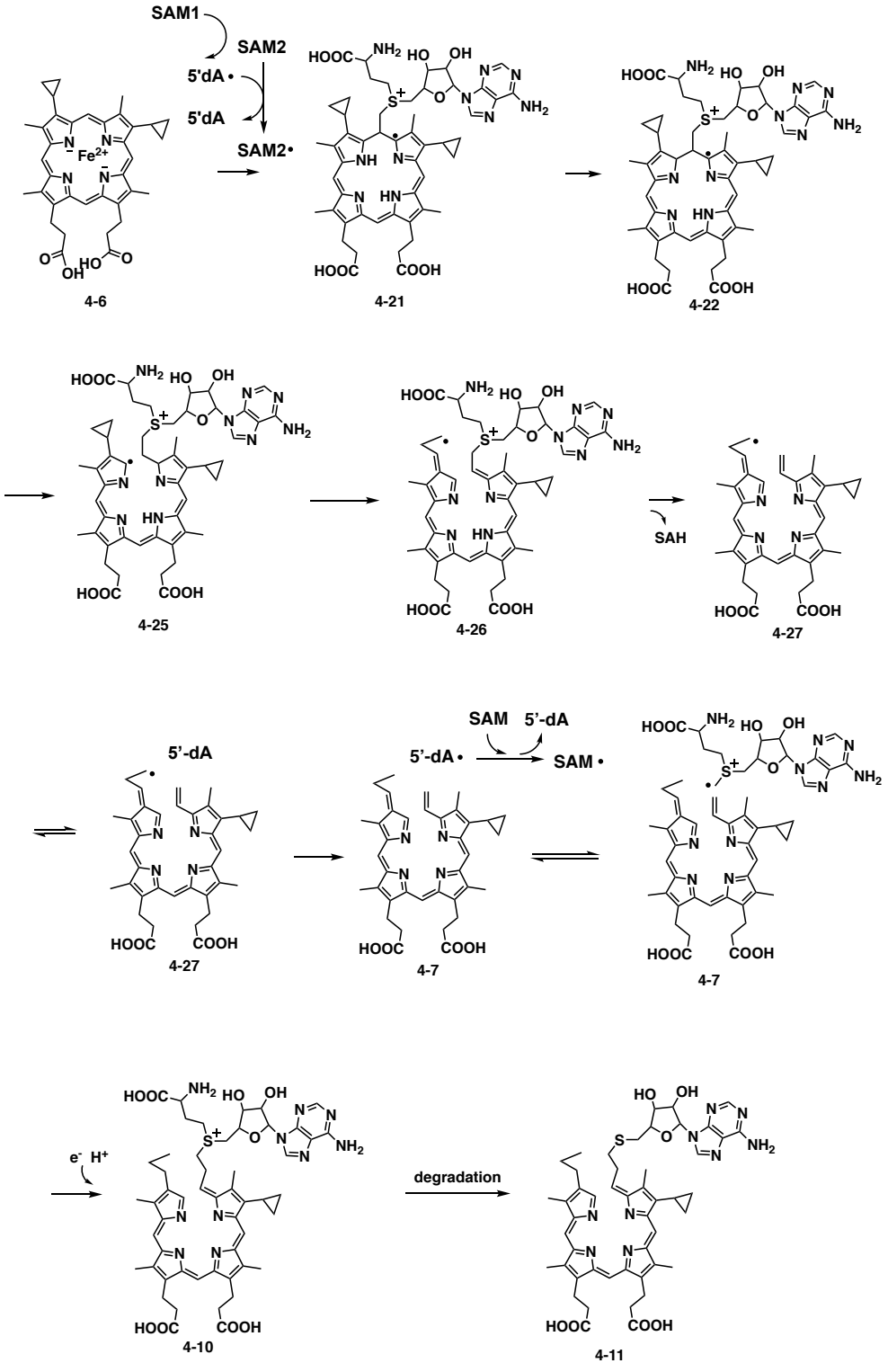


Figure 4.35 Proposed mechanism for the formation of the crosslinking species in ChuW catalyzed cyclopropyl heme degradation.

4.2.8. Proposed mechanisms for the ring opening product of ChuW catalyzed heme degradation

Based on the previous results, cyclopropyl substrate analog studies and the proposed mechanism for NosN[37-38], TbtI[42], C10P[41], a similar mechanism can be proposed for the heme degradation by ChuW (Figure 4.36).

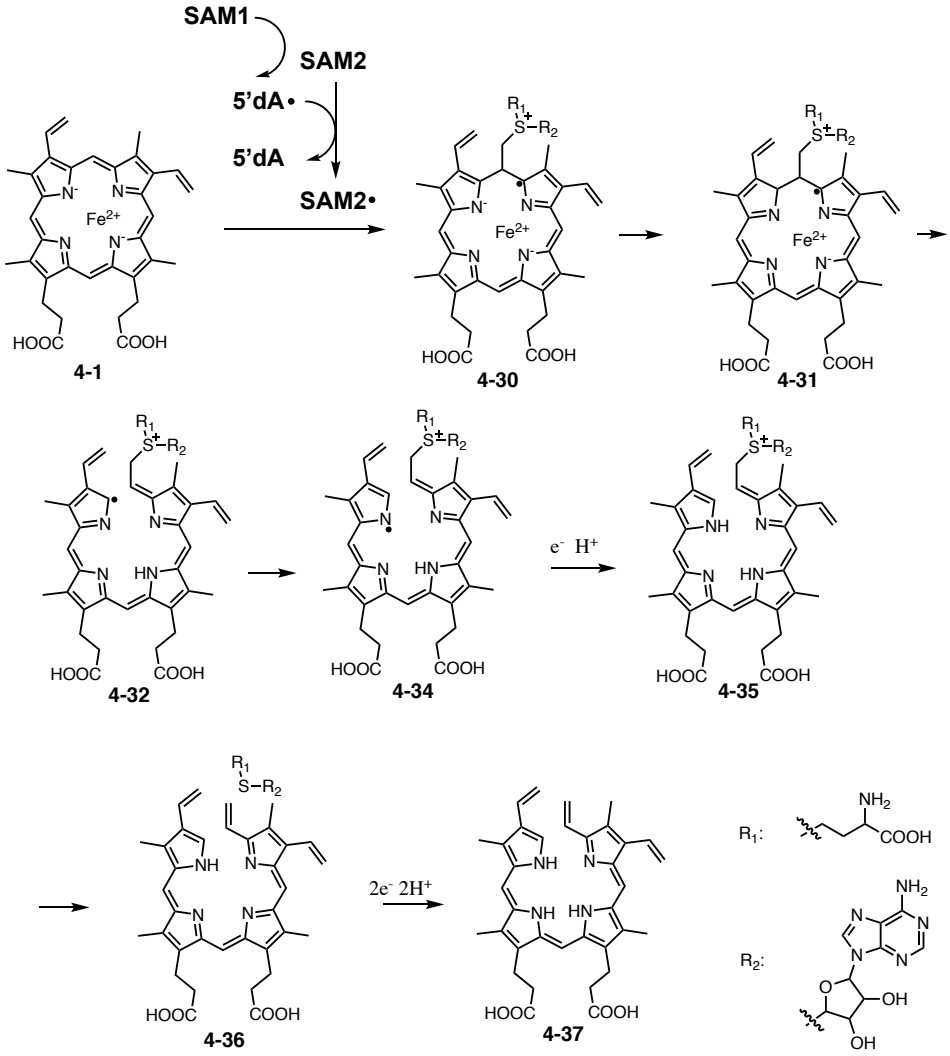


Figure 4.36 Proposed mechanism of ChuW catalyzed heme degradation.

4.3. Conclusion

Cyclopropyl heme analog was designed as a radical probe for ChuW catalyzed heme degradation. ChuW can take cyclopropyl heme analog as a substrate and several ring opening products were detected by LC-MS/MS, HPLC. A protoporphyrin ring opening product was further characterized by hydrogenation and is considered more stable for NMR characterization. However, there is no NMR spectrum currently available for this porphyrin compound, because the low yield of the enzymatic product and the potential light and air sensitivity of the compound.

Similar as anaerobilin, porphyrin cleavage products are photolabile, including SAM anaerobilin adducts. The adducts were further confirmed by isotope labeling SAM studies. The cyclopropyl ring opening products and the SAM adducts suggested that a radical is generated on the porphyrin ring of heme by the attack of a SAM radical species. Based on the result, a mechanism was deduced for heme degradation by ChuW.

4.4. Experimental methods

4.4.1. Synthesis of cyclopropyl heme analog

The overall scheme is shown in Figure 4.5. Each detailed step is shown as below:

Synthesis of the protoporphyrin-IX dimethyl ester (4-3)

The protoporphyrin-IX (100 mg) was added to a stirred solution of the trimethyl orthoformate (10 ml) and the methanol (10 ml). Concentrated sulfuric acid (1 ml) was then added slowly, with cooling to room temperature, and the solution was stirred at room temperature for 45 minutes. The mixture was then diluted with water, the pH was adjusted to 4.5 by the addition of dilute sodium hydroxide and the mixture was extracted

with dichloromethane. The organic extracts were washed with water and the solvent was removed under reduced pressure. The resulting solid was purified by column chromatography (silica gel, 20:1 CH_2Cl_2: MeOH) to give protoporphyrin-IX dimethyl ester. ^{1}H NMR (400 MHz, Chloroform-d) δ 10.09 (d, J = 4.2 Hz, 2H), 9.97 (s, 2H), 8.24 (m, 2H), 6.36 (d, J = 17.8 Hz, 2H), 6.19 (d, J = 11.5 Hz, 2H), 4.39 (t, J = 7.8 Hz, 4H), 3.82 – 3.53 (m, 18H), 3.28 (t, J = 7.8 Hz, 4H).

Synthesis of 3,8-di- [3' (R, S) -(1' -pyrazolinyl)]-deuteroporphyrin IX dimethyl ester (4-4)

All reactions were carried out in the dark. A protective shield was used during diazomethane cycloaddition. The starting protoporphyrin-IX dimethyl ester (100 mg) was dissolved in dichloromethane (100 ml) and an ethereal solution of diazomethane (50 ml) was added to the reaction mixture. The resulting solution was sealed with a rubber septum and kept at room temperature for 48 – 96 h. The solvent was degassed by bubbling with nitrogen and the resulting solution evaporated. The resulting solid was purified by column chromatography (silica gel, 100:1 - 10:1 CH_2Cl_2: acetone) to give 3,8-di- [3' (R,S)-(1' -pyrazolinyl)]-deuteroporphyrin IX dimethyl ester. ^{1}H NMR (400 MHz, Chloroform-d) δ 10.05 (d, J = 21.4 Hz, 2H), 9.78 (dd, J = 34.3, 6.0 Hz, 2H), 7.03 – 6.83 (m, 2H), 5.64 – 5.49 (m, 2H), 4.94 – 4.75 (m, 2H), 4.39 (td, J = 7.8, 3.2 Hz, 4H), 3.74 – 3.50 (m, 18H), 3.29 (td, J = 7.8, 2.7 Hz, 4H), 2.94 – 2.78 (m, 2H), 2.32 – 2.16 (m, 2H).

Synthesis of 3,8-dicyclopropyldeuteroporphyrin IX dimethyl ester (4-5)

The 8-di- [3' (R, S) -(1' -pyrazolinyl)]-deuteroporphyrin IX dimethyl ester (31 mg) were dissolved in o-dichlorobenzene (20 ml) and heated at reflux for 45 min. The reaction solution was diluted with hexane (200 ml), and the formed residue was filtered and purified by column chromatography (silica gel, 100:1 CH_2Cl_2: acetone) to give the 3,8-dicyclopropyldeuteroporphyrin IX dimethyl ester. ^{1}H NMR (400 MHz, Chloroform-d) δ 10.41 (d, J = 8.6 Hz, 2H), 10.03 (d, J = 10.0 Hz, 2H), 4.43 (m, 4H), 3.80 – 3.55 (m, 18H), 3.32 (m, 4H), 3.14 – 3.01 (m, 2H), 1.70 (m, 4H), 1.54 – 1.44 (m, 4H).

Synthesis of Ferric 3,8-dicyclopropyldeuteroporphyrin IX (4-6)

Synthesis of ferric 3,8-dicyclopropyldeuteroporphyrin IX was followed a protocol to synthesize ferric meso-methylmesoheme regioisomers.[84] The 3,8-dicyclopropyldeuteroporphyrin IX dimethyl ester was hydrolyzed in 25% (v/v) HCl at 25 °C in the dark for 8 h. The hydrolysis reaction was worked up by neutralizing to pH 2-4, extracting with CH_2Cl_2, washing the organic phase with water before drying it with anhydrous Na_2SO_4, and removing the solvent under vacuum. The porphyrins were used without further purification for the subsequent reaction in which the iron was inserted by the ferrous sulfate method. An argon-purged, saturated aqueous Fe_2SO_4 solution (1mL) was added to a stirred, argon-purged solution of the porphyrins in 1 mL of pyridine and 20 mL of acetic acid. The mixture was warmed for 10 min under a stream of argon. The stirred reaction mixture was then exposed to the air while it cooled to promote autooxidative formation of the ferric state before it was combined with 25 mL of brine and extracted with diethyl ether. The organic layer was washed with 25% HCl to remove

unreacted porphyrin, washed with water, and dried over Na_2SO_4. Removal of the solvent under vacuum provided the ferric 3,8-dicyclopropyldeuteroporphyrin IX, which were purified by reversed-phase HPLC. HRMS m/z 644.2113, calculated for $FeC_{36}H_{36}N_4O_4^+$ 644.2080. λ_{max} (CH_2C1_2) 420, 540, 640 nm.

4.4.2. Overexpression and purification of ChuW

chuW was cloned into pTHT vector (modified from pET28b with TEV cleavage site). ChuW was over expressed in *E. coli* strain BL21 DE3 and was co-expressed with a plasmid encoding the *suf* operon. A 15 mL overnight culture was grown in LB medium in the presence of 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. This was then added to 1.5 L of LB medium containing 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. The cultures were incubated at 37°C with shaking (180 rpm) until the OD600 reached 0.45. The culture was then incubated at 4°C without shaking for 1 hour. Expression was induced by the addition of 100 μM IPTG followed by ferrous ammonium sulfate (120 mg) and cysteine (120 mg) and the culture was grown at 15 °C with shaking (110 rpm) overnight. The cells were harvested by centrifugation and stored in liquid nitrogen until further use. All steps for protein purification were carried out in an anaerobic chamber (COY laboratories). Cell pellets (~12 g) were thawed and resuspended in 30 mL lysis buffer (100 mM Tris-HCl, pH 7.5) in the presence of lysozyme (0.2 mg/mL) and benzonase (100 units). This mixture was cooled in an ice-bath with continuous stirring for 1-2 hrs. The suspension was then sonicated to complete cell lysis for 5 × 120 s using a Misonix Sonicator XL-2000 (setting = 9). Cell debris was removed by centrifugation, and the lysate was loaded onto a Histrap column pre-

equilibrated with lysis buffer. The column was washed with 5-10 column volumes of wash buffer (100 mM Tris-HCl, 30 mM Imidazole, pH 7.5). The CysS protein was then eluted using elution buffer (100 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH 7.5). The brown-colored fractions, containing the desired protein, were pooled and buffer exchanged to final buffer (100 mM Tris-HCl, pH 7.5 at 25 °C, 30% glycerol) using Econo-Pac 10DG size exclusion chromatography. The purified CysS was divided into aliquots in eppendorf tubes, frozen in liquid nitrogen, and stored at −80 °C, submerged in liquid nitrogen, until future use. Each tube has a pin hole in the cap to allow for gas release during thawing. Protein concentration was measured by the absorbance at 280 nm (A280) when the protein was exposed to air or the $[4Fe\text{-}4S]^{2+}$ cluster was reduced to $[4Fe\text{-}4S]^{1+}$ by dithionite (prevent the interference from absorbance from $[4Fe\text{-}4S]^{2+}$) with an extinction coefficient calculated using the ProtParam tool of the ExPASy proteomics server.

4.4.3. *In vitro* activity assay with ChuW:

For the UV-Visible spectroscopy of monitoring ChuW activity, ChuW (5 μM) was incubated with 20 μM substrate and EcFld (5 μM) /EcFpr (2 μM), NADPH (200 μM) reducing system anaerobically in dark. Assays were initiated by the addition of SAM (250 μM) and UV spectra were recorded every 10-20 minutes.

For the HPLC and LCMS analysis of ChuW activity, 150 μM ChuW was incubated with 300 μM substrate, 1 mM SAM, 1 mM NADPH, 100 μM EcFld, and 50 μM EcFpr in pH 7.5, 100 mM phosphate buffer anaerobically for 2.5 h at room temperature and in the dark. The assay mixture was applied to the SPE column pre-

equilibrated with 1 eq of methanol, 1 eq of H_2O, and 1 eq of assay buffer. Once the assay mixture ran through the column, it was washed with 1eq of buffer,2 eq of water,1eq of 1%TFA, and 3eq of water, before elution in 100% methanol in dark. The elution solution was further lyophilized before analyzed by HPLC and LCMS.

4.4.4. HPLC parameters for the analysis of the ChuW reaction mixture

HPLC analysis was performed on a ZORBAX Eclipse XDB-C18 column (15cm 4.6mm, 5μm particles, Agilent Technologies, flow rate1⁄41mL/min). Typical injection volumes were in the range of 10–80μL. Data were processed using ChemStation ver. B.04.01 SP1 (Agilent Technologies).

LC conditions: A- 0.1 % formic acid in H_2O, B- 0.1 % formic acid in acetonitrile

LC method: 0 min – 100% A, 2 min – 100% A, 4 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

4.4.5. LC-MS parameters for the analysis of the ChuW reaction mixture

LC-ESI-TOF-MS on small molecules was performed using an Agilent 1260 HPLC system equipped with a binary pump and a 1200 series diode array detector followed by a MicroToF-Q II mass spectrometer (Bruker Daltonics) using an ESI source in positive mode. The analysis was performed on a ZORBAX Eclipse XDB-C18 column (15 cm x 4.6 mm, 5 μm particles, Agilent).

LC conditions: A- 5 mM Ammonium acetate buffer, pH 6.6, B- 75 % Methanol and 25 % Water.

LC method: 0 min – 100% A, 2 min – 100% A, 4 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

5. TRAPPING A FOUR-CARBON FRAGMENT IN THIC CATALYZED REARRANGEMENT IN THIAMINE BIOSYNTHESIS

5.1. Introduction

Thiamine pyrophosphate (vitamin B_1) is a crucial cofactor that is present in all living systems, in which it helps to catalyze essential biochemical reactions. Thiamine pyrophosphate is synthesized in the cytosol and is required in the cytosol for the activity of transketolase and in the mitochondria for the activities of pyruvate-, oxoglutarate-, branched chain keto acid dehydrogenases and thiamin acts as a stabilizer of an acyl carbanion biosynthon. The deficiency of thiamine in the diet will lead to the peripheral nervous system disease beriberi in human.

In all organisms, the thiazole and pyrimidine moieties of thiamin monophosphate (ThMP) are generated in separate branches of the pathway and then joined by a coupling enzyme. The mechanisms of most enzymes in thiamin biosynthesis pathway are well understood.[85-86] One of the exceptions is ThiC, a pyrimidine synthase in the thiamin biosynthetic pathways of bacteria and plants.

ThiC catalyzes the complex rearrangement of aminoimidazole ribonucleotide (AIR) **5-1** to 4- amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) **5-2** (Figure 5.1). The origin of all the atoms of the product and the fate of most of the atoms of the substrate have been determined using isotope labeling studies.[4, 15, 87-88]

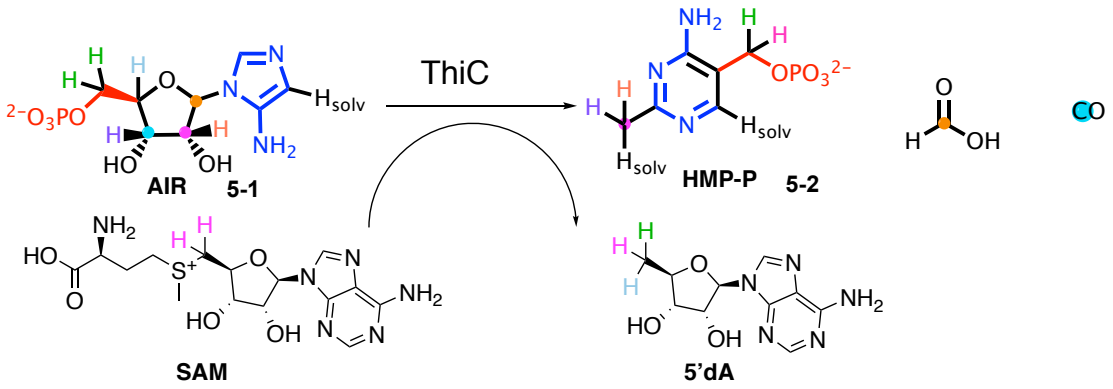


Figure 5.1 ThiC catalyzed rearrangement from AIR to HMP-P

From the experimental data, we know that the proS hydrogen at C5′ is first abstracted by 5′-dA radical and then later replaced by a hydrogen atom from 5′-dA with retention of stereochemistry. The resulting 5′-dA radical has the second abstraction of a hydrogen atom from C4′ of AIR. C2′ of AIR becomes the methyl group C8 of HMP-P and C4′, C5′ remain in the HMP-P. C1′ of AIR is released as formate and C3′ is lost as carbon monoxide. N5 of AIR is retained as the N4 of the HMP-P. C3′ hydrogen of AIR is transferred to C8 of HMP-P while C2′ hydrogen remains in the C8 of HMP-P. One hydrogen on C8 and the hydrogen on C6 of HMP-P are from the solvent.

Based on the current data, a mechanism for ThiC catalyzed rearrangement can be proposed (Figure 5.2). The 5′-dA radical first abstracts a hydrogen atom from C5′ of AIR. The resulting radical goes through a beta scission and then a reduction by the amino imidazole to give **5-5**. The radical on the amino imidazole induces the cleavage of the N-glycosyl bond. Then the amino imidazole is reoriented in the active site to add on the enol phosphate to form **5-8**, which abstracts a hydrogen atom from 5′-dA. The 5′-dA

radical further abstracts a hydrogen atom from C4′. Cyclization and hydrogen atom transfer of the resulting radical give **5-13**. Then, C1′ leaves as formate to form the radical stabilized carbanion **5-14**. After protonation, the C3′ radical goes through a ribonucleotide reductase type of dehydration to form α-keto radical **5-16**. Similar as the Dowd–Beckwith ring-expansion reaction, another cyclization gives bicyclic intermediate **5-17** followed by two beta scission and oxidation to give acylium ion **5-19**. CO elimination and tautomerization give the final product HMP-P **5-2**.

A ThiC paralog, BzaF, can catalyze the conversion of AIR to hydroxybenzimidazole – one form of the axial ligands of vitamin B_{12}. Labeling studies have established that the early steps in the ThiC and BzaF-catalyzed reactions are same. EPR studies on BzaF using AIR isotopologues show a transient radical with a structure that is consistent with intermediate **5-5**.[89] Since desamino-AIR is not a substrate for ThiC, the amino group may play a role for in activating the substrate for the electron transfer shown in the conversion of **5-4** to **5-5** in Figure 5.2.

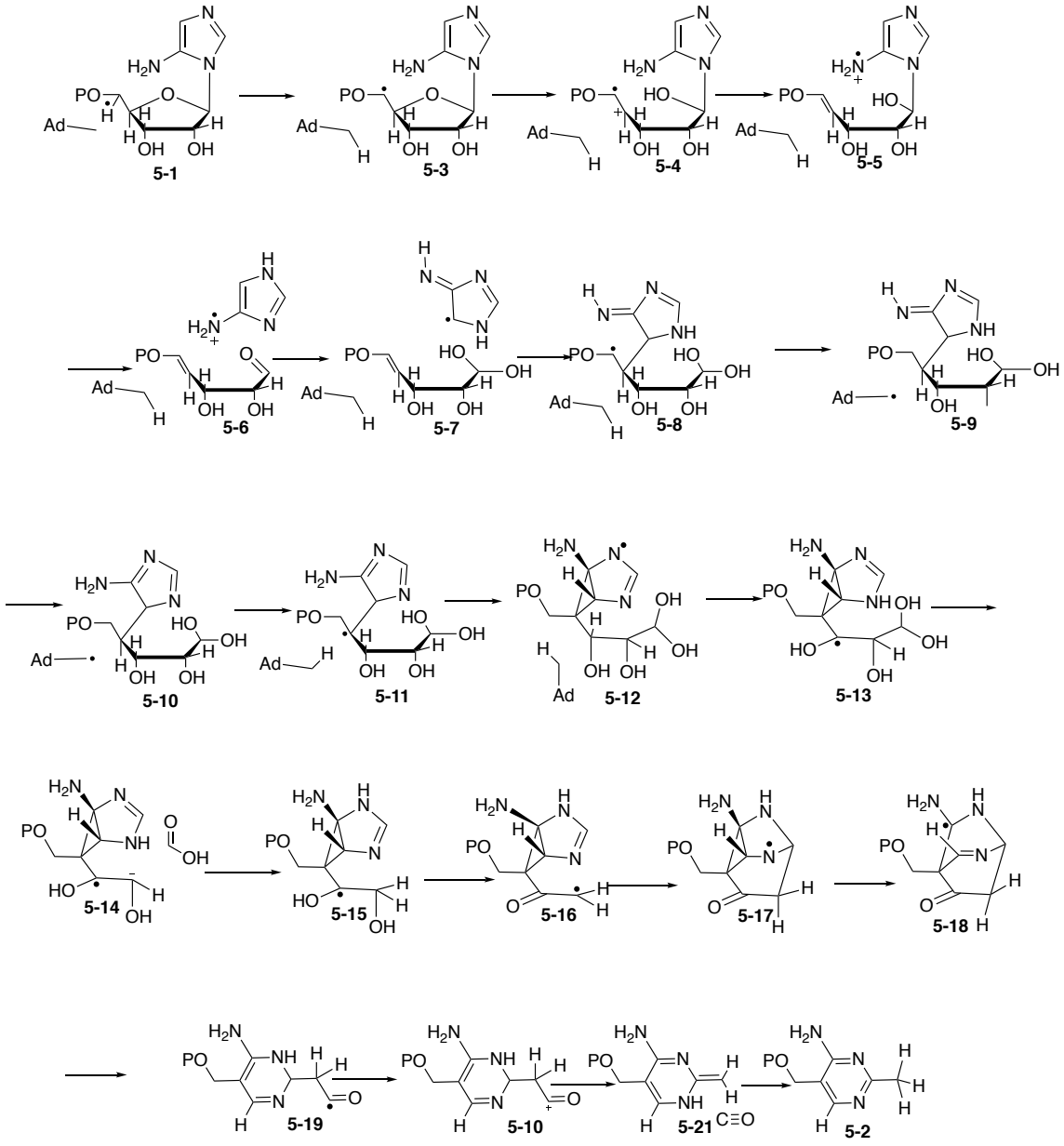


Figure 5.2 Proposed mechanism for ThiC catalyzed AIR **5-1** to HMP-P **5-2** rearrangement.

Compared with the canonical CX_3CX_2C cluster-binding motif for radical SAM enzyme, ThiC adopts an unusual CX_2CX_4C motif for the iron-sulfur cluster binding. Several crystal structures of ThiC with bound [4Fe-4S] cluster, small molecules (SAH,

5′-dA, methionine, AIR, imidazole ribonucleotide) are available. These structures not only reveal a novel second iron site involved in amino acid binding but also give an insight into possible protein residues in the active site that may involve in the catalysis. The active D383N mutant suggests that the N-glycosyl bond cleavage is not strongly dependent on acid catalysis. Several other mutants (D383N, M248L, C474S and H417A) still needs further investigation.

Although the activity of ThiC was reconstituted *in vitro* more than ten years ago,[87] we still don't know most of the steps in the mechanism because of the low activity and narrow range of substrates tolerance. While we tried to identify the new products and metabolites in ThiC mutants, a four-carbon fragment from AIR was detected during ThiC catalyzed reaction.

5.2. Result and discussion

5.2.1. Identification of bound metabolites co-purifying with ThiC

When wild-type ThiC (wtThiC), recombinantly over-expressed in *E. coli* in LB medium and purified in a glove box, several bound metabolites were detected by HPLC under UV of 250 nm after ThiC was heat denatured (Figure 5.3). By comparing with the UV and the retention time of the standard, around 4 min peak is identified as SAM and around 17 min peak as 5′-dA. After the treatment of alkaline phosphatase from calf intestinal (CIP), the early retention time peaks were disappeared or decreased with the increase of 16 min peak and appearance of a new peak around 15 min. For the ThiC mutant C474S, besides the same peaks that were detected in the wtThiC, there was one

more peak around 2 min, which disappeared under CIP treatment with the formation of a new peak around 14 min (Figure 5.4).

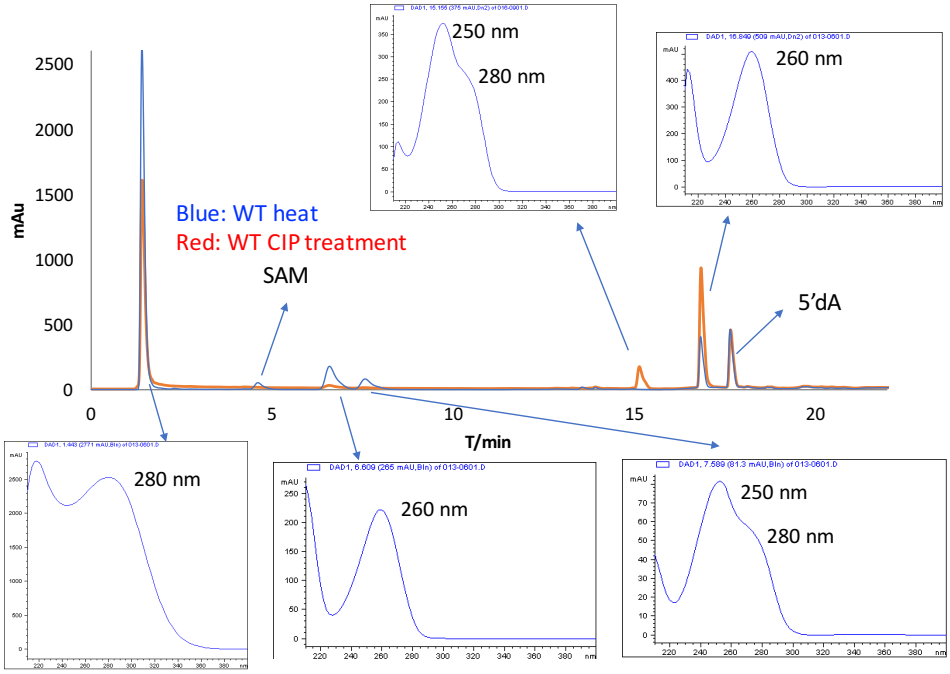


Figure 5.3 HPLC analysis of bound metabolites in ThiC under 250 nm. UV spectrum of different bound metabolites metabolites.

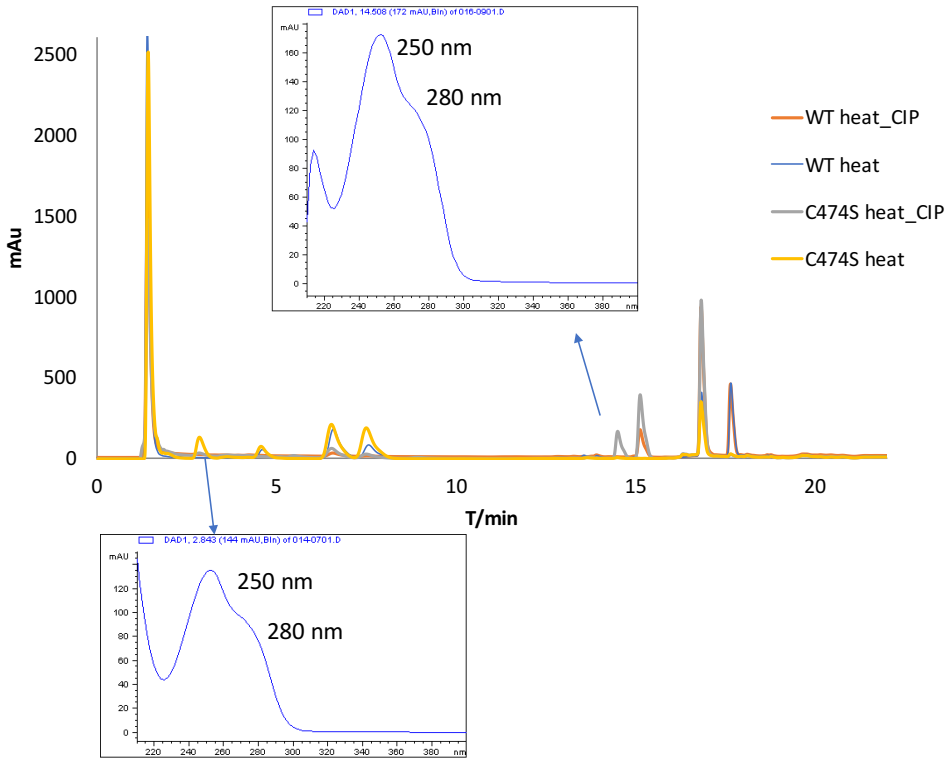


Figure 5.4 HPLC analysis of bound metabolites in the ThiC mutant C474S under 250 nm. UV spectrum of different bound metabolites metabolites.

Based on the UV similarity and co-elution studies of the bound metabolites with different nucleosides, the 17 min peak is identified as adenosine (Figure 5.5). The 15 min peak is identified as deoxyguanosine. The 14 min peak is identified as guanosine. So, the earlier time point peaks on the HPLC could probably be the corresponding nucleotides.

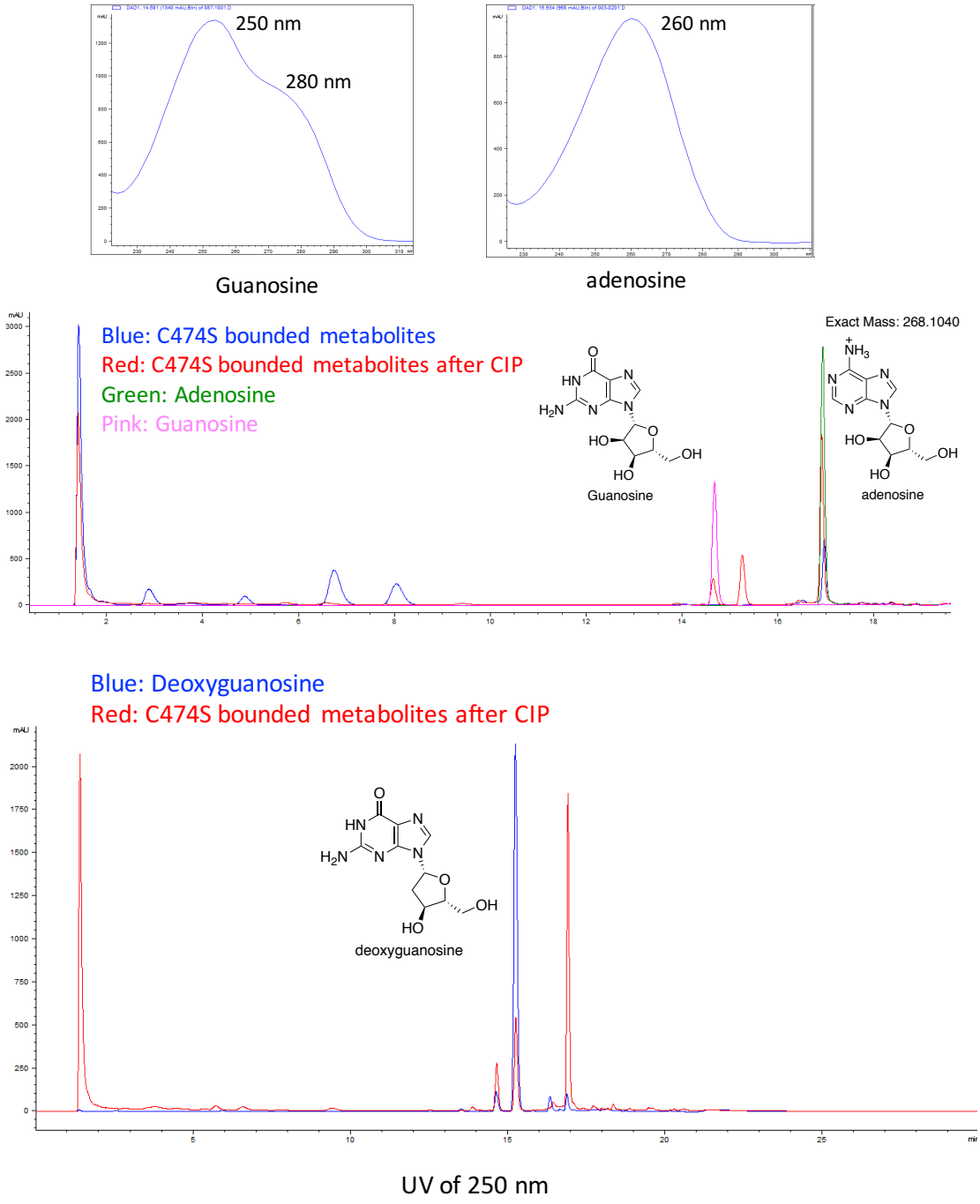


Figure 5.5 Upper panel: UV spectrum of guanosine and adenosine. Middle panel: co-elution HPLC traces of bound metabolites in the ThiC mutant C474S and guanosine, adenosine under 250 nm. Lower panel: co-elution HPLC traces of bound metabolites in the ThiC mutant C474S and deoxyguanosine.

5.2.2. Detection of a four-carbon fragment in ThiC reaction by LC-MS/MS

When *Arabidopsis thaliana* ThiC (AtThiC) was incubated with AIR, SAM, dithionite in an anaerobic condition, a new peak with a mass of 204.10 was detected by LC-MS/MS after heat quenching (Figure 5.6). This peak only presents in the full reaction and is absent in all other controls. By using isotope labeling AIR substrates, it was found that four carbons from the ribose part of the AIR but not the nitrogen of the amino group and C1′ of AIR were incorporated in this new compound (Figure 5.7).

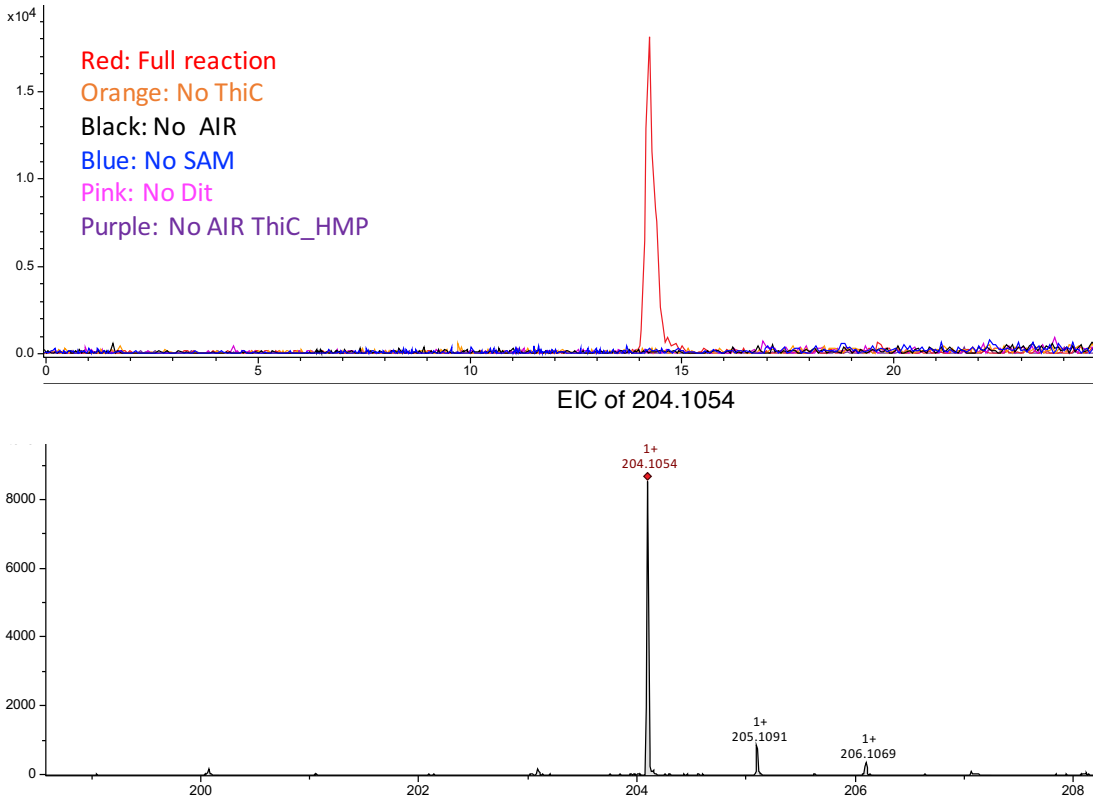


Figure 5.6 LC-MS analysis of ThiC catalyzed AIR to HMP rearrangement. Upper panel: EICs of 5′-dA $[M+H]^+$ (204.10±0.02). Lower panel: mass spectrum of 204 compound at 14 min.

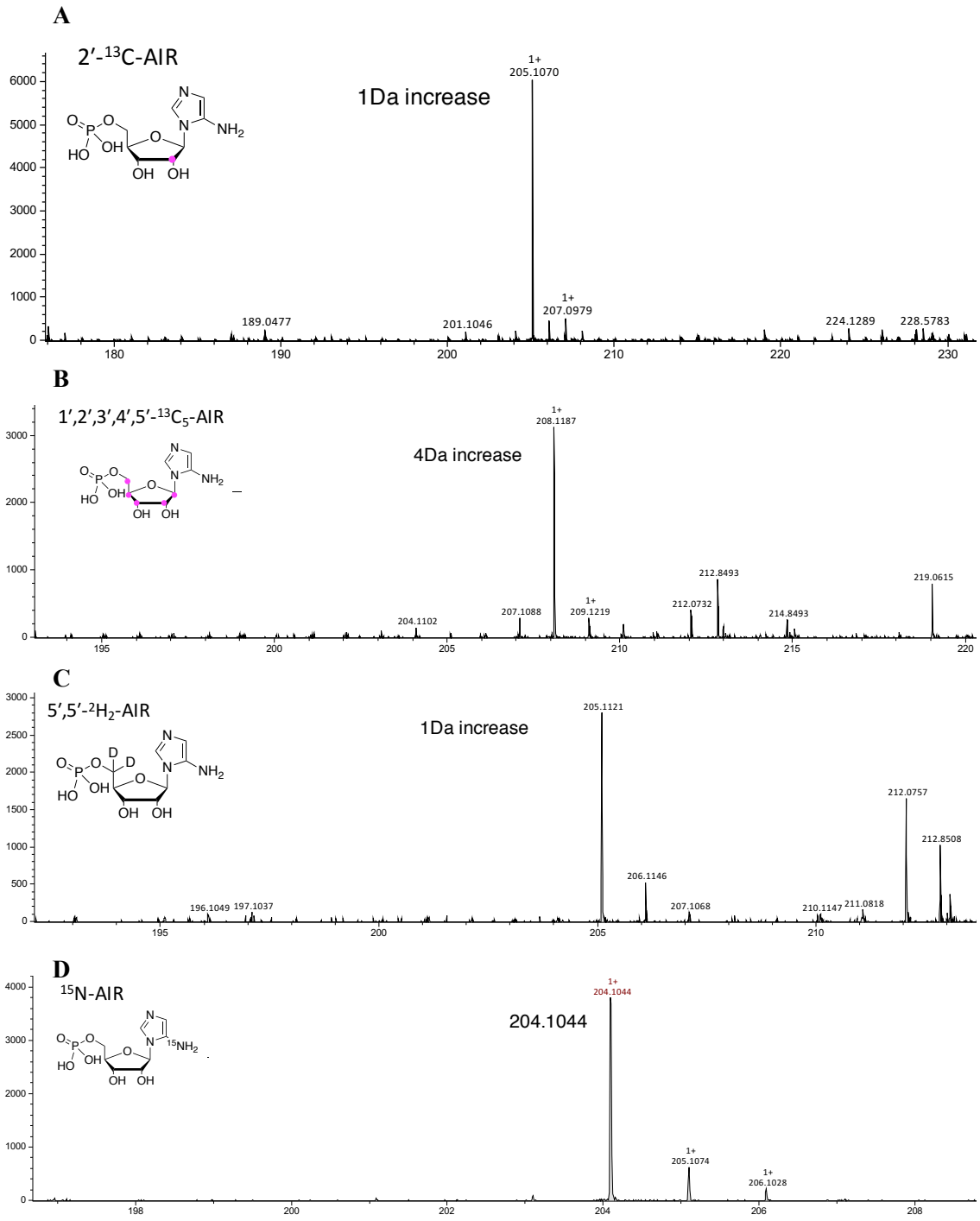


Figure 5.7 Isotope labeling pattern in 204.10 shunt product in ThiC catalyzed reaction. Mass spectrum of the in 204.10 compound when the substrate is (A) 2′-^{13}C-AIR; (B) 1′,2′,3′,4′,5′-$^{13}C_5$-AIR; (C) 5′,5′-$^{2}H_2$-AIR; (D) ^{15}N-AIR; (E) 1′-^{13}C-AIR; (F) 2′-^{2}H-AIR; (G) 4′-^{2}H-AIR.

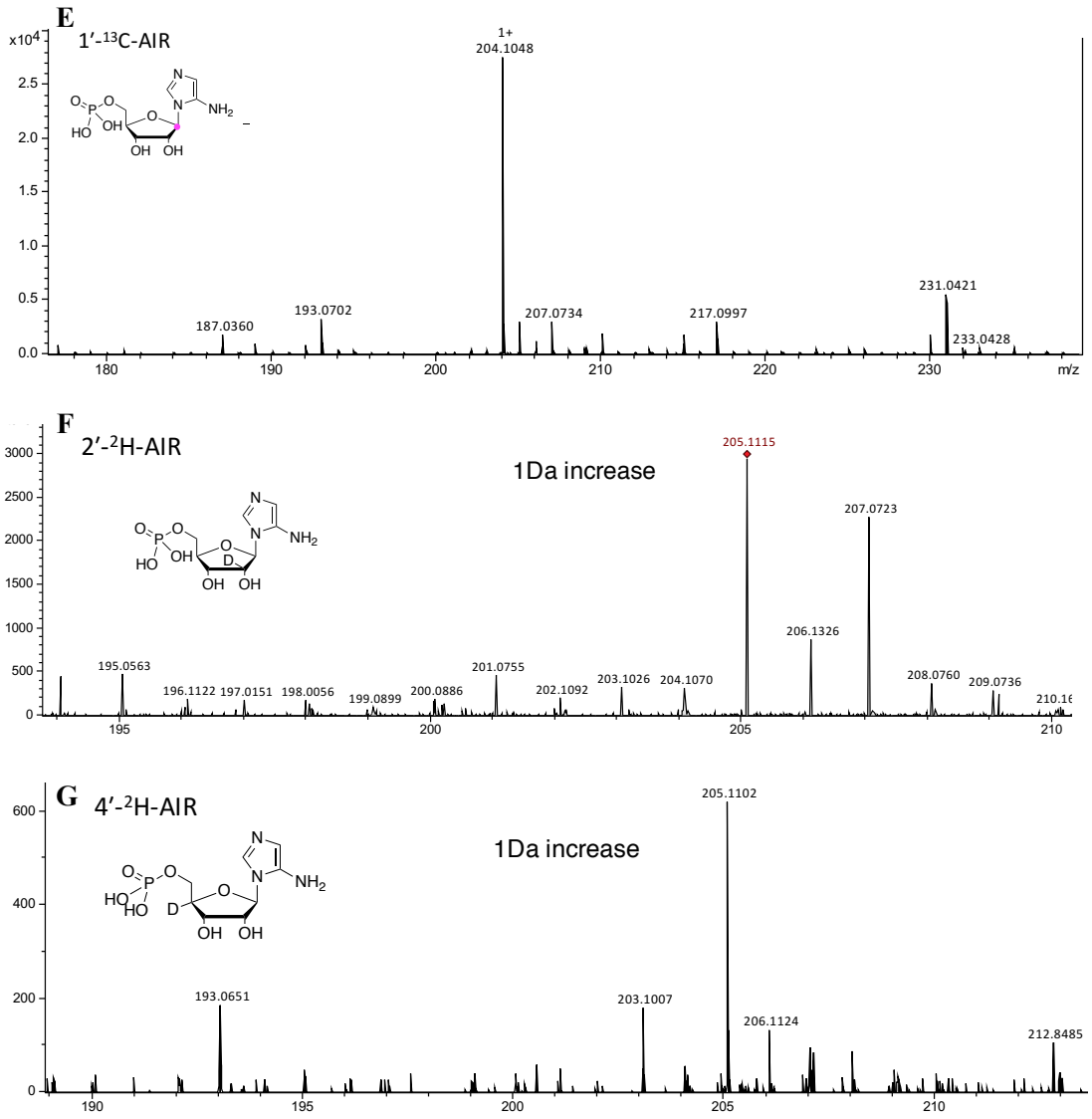


Figure 5.7 Isotope labeling pattern in 204.10 shunt product in ThiC catalyzed reaction. Mass spectrum of the in 204.10 compound when the substrate is (A) 2′-^{13}C-AIR; (B) 1′,2′,3′,4′,5′-$^{13}C_5$-AIR; (C) 5′,5′-$^{2}H_2$-AIR; (D) ^{15}N-AIR; (E) 1′-^{13}C-AIR; (F) 2′-^{2}H-AIR; (G) 4′-^{2}H-AIR.

5.2.3. Identification of the methionine as the trapping agent for the four-carbon fragment of the AIR

When CD_3-SAM was used in the reaction, there was an increase of three mass units in the mass of 204.10 peak. Similarly, when $^{13}CH_3$-SAM was used as the cofactor, there was an increase of one mass in the mass of 204.10 peak (Figure 5.8). Considered the error range of the mass instrument we used, the error of the measured mass and the exact mass should be less than 10 ppm. Based on the reactants in ThiC reaction, one of the most possible chemical formula of 204.1048 contains a sulfur atom. These results lead us to think the sulfur as well as the methyl group of SAM could be incorporated into this shunt product of the ThiC reaction.

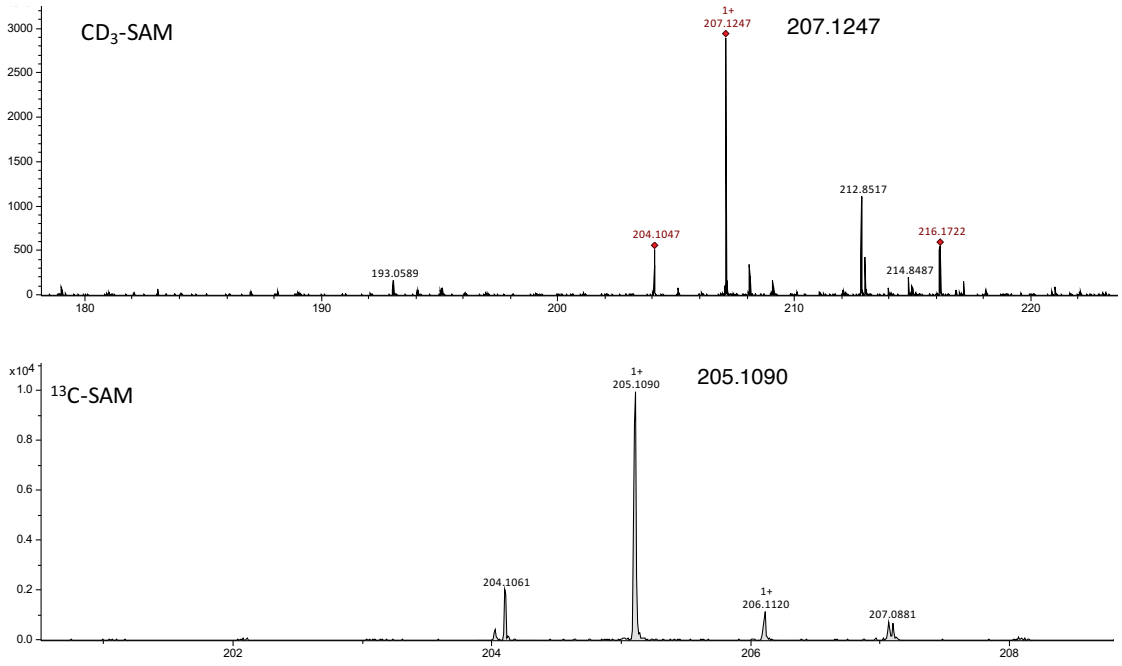


Figure 5.8 Isotope labeling pattern in 204.10 compound from different isotope labeling SAM. Mass spectrum of the in 204.10 compound when (Upper panel) CD_3-SAM; (Lower panel) $^{13}CH_3$-SAM.

In order to explore whether the other parts of SAM were incorporated in this 204.10 peak, different labeling SAM was biosynthesized. When (adenosyl-$^{13}C_{10}$)-SAM was used as the cofactor in ThiC reaction, the mass of byproduct peak remained the same, which suggested that the adenosyl part of SAM was not incorporated into this compound (Figure 5.9). However, when (methionine-D_8)-SAM was used in the reaction, there was an increase of eight mass units in this compound (Figure 5.9).

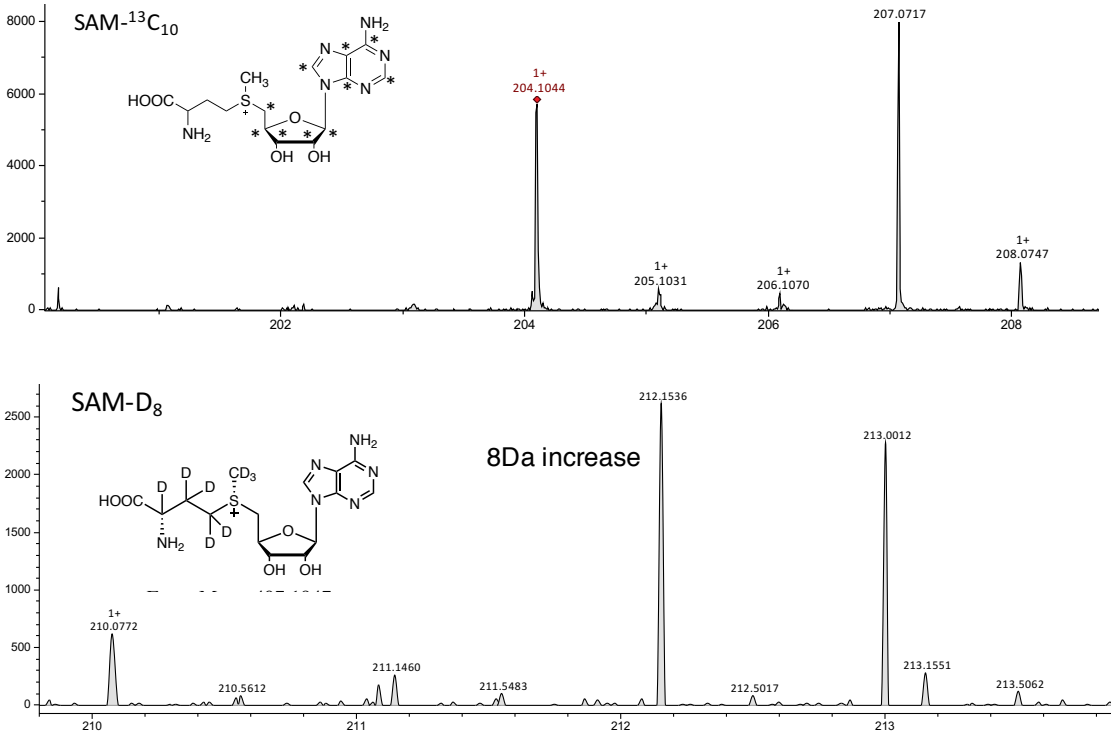


Figure 5.9 Isotope labeling pattern in 204.10 compound from different isotope labeling SAM. Mass spectrum of the in 204.10 compound when different isotope labeling SAM was used in the ThiC catalyzed reaction. Upper panel: (adenosyl-$^{13}C_{10}$)-SAM; Lower panel: (methionine-D_8)-SAM.

To further explore whether the methionine in the 204.10 peak comes from SAM, fivefold excess of D_8-methionine (compared with the concentration of SAM) was incubated in the ThiC reaction. LC-MS analysis of the reaction mixture revealed that more than 80% of the product was deuterium labeled (212.15) (Figure 5.10). This suggests that methionine part that gets incorporated into the peak does not have to come from the cleavage product of the SAM.

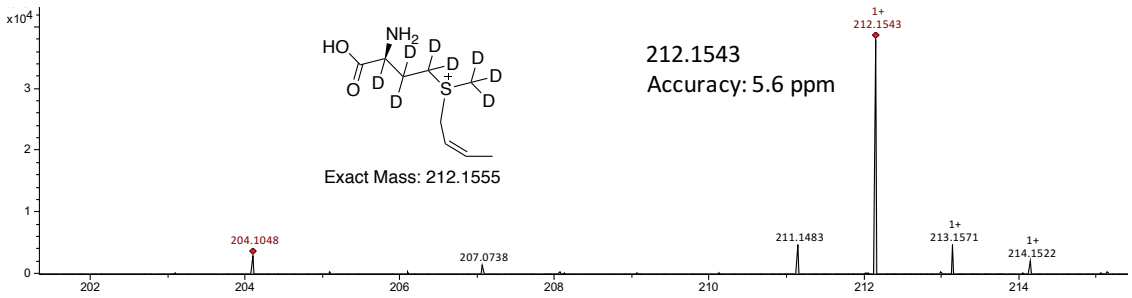


Figure 5.10 Isotope labeling pattern in 204.10 compound when excess of D_8-methionine was added to the ThiC catalyzed reaction.

In order to probe whether the AIR fragments trapped by methionine happened in the enzyme active site, D-methionine, L-methionine, cysteine, L-homocysteine were also incubated with the ThiC reaction respectively. Only D/L-methionine showed the incorporation in the unknown compound, while there is no much difference between D-methionine or L-methionine.

5.2.4. Determination of possible reference compounds for the detected shunt product in ThiC catalyzed reaction

Based on the experimental data in section 5.2.2-5.2.4, the shunt product consists of four carbon atoms (2′,3′,4′,5′) from AIR substrate and methionine from SAM or buffer. It's not known how the four-carbon fragment is connected to the methionine. Based on the ThiC proposed mechanism, the crystal structure and the electrophilic property of the atoms in methionine, the most possible structure of the shunt product is the sulfonium compounds **5-21**, where the four-carbon is linked to methionine through the sulfur atom (Figure 5.11). In one of the conformation in ThiC, the distance between sulfur atom of SAH and the 5′C of the AIR is 4.4 Å (Figure 5.12). Once the ribose ring

of AIR is broken, the distance could be smaller. Synthesis of compounds **5-21** are in progress.

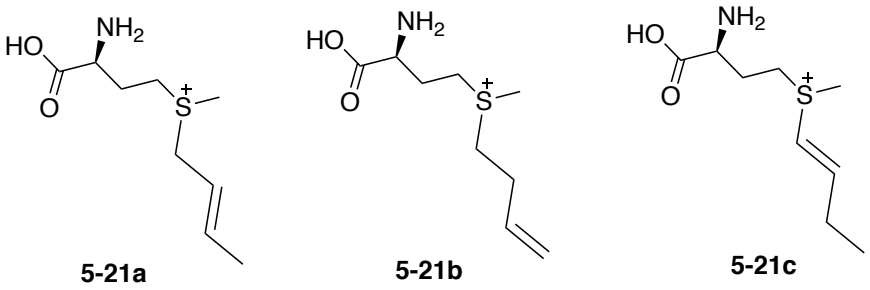


Figure 5.11 Possible structures for the shunt product.

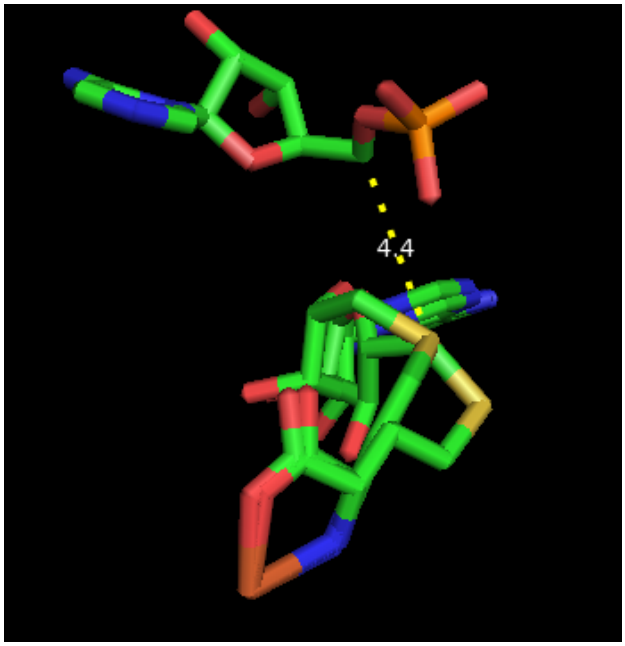


Figure 5.12 The distance between sulfur atom of SAH and the 5′C of the AIR in ThiC crystal structure (4S28)

Based on the analysis of MS/MS fragments of the shunt product, it is also possible that the four-carbon is linked to methionine through the nitrogen atom of the methionine (Figure 5.13, 5.15). Compounds **5-22a** and **5-22b** were synthesized. However, from the retention time and MS/MS fragments of these compounds, neither of them match the enzymatic shunt product generated in the ThiC catalyzed reaction. It's possible that **5-21** can have an arrangement to form **5-22**, which gives the observed MS/MS fragmentations.

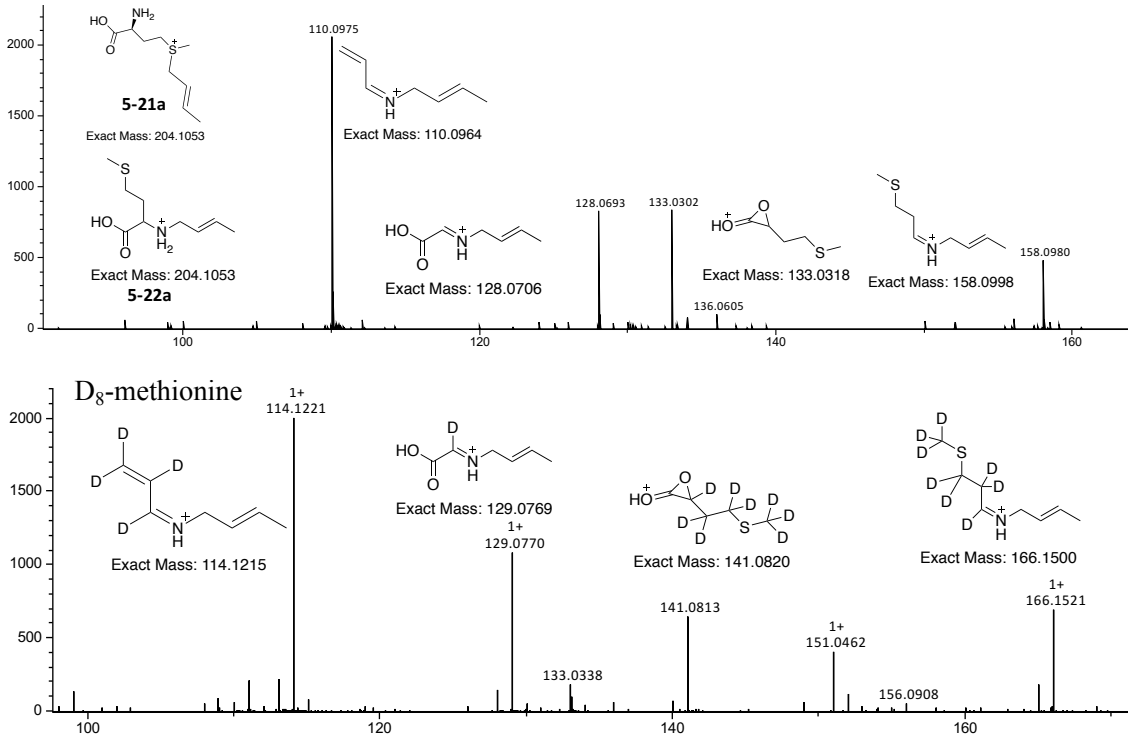


Figure 5.13 Mass fragmentation of the shunt product (204.1053) when ThiC catalyzed reaction with SAM, AIR, dithionite and methionine (upper panel), D_8-methionine (lower panel).

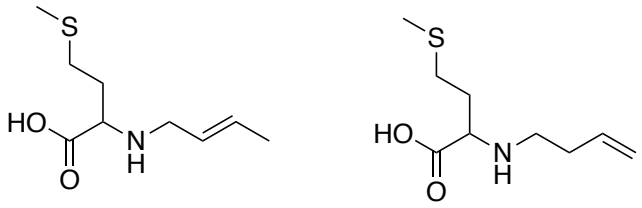


Figure 5.14 Possible structures for the shunt product.

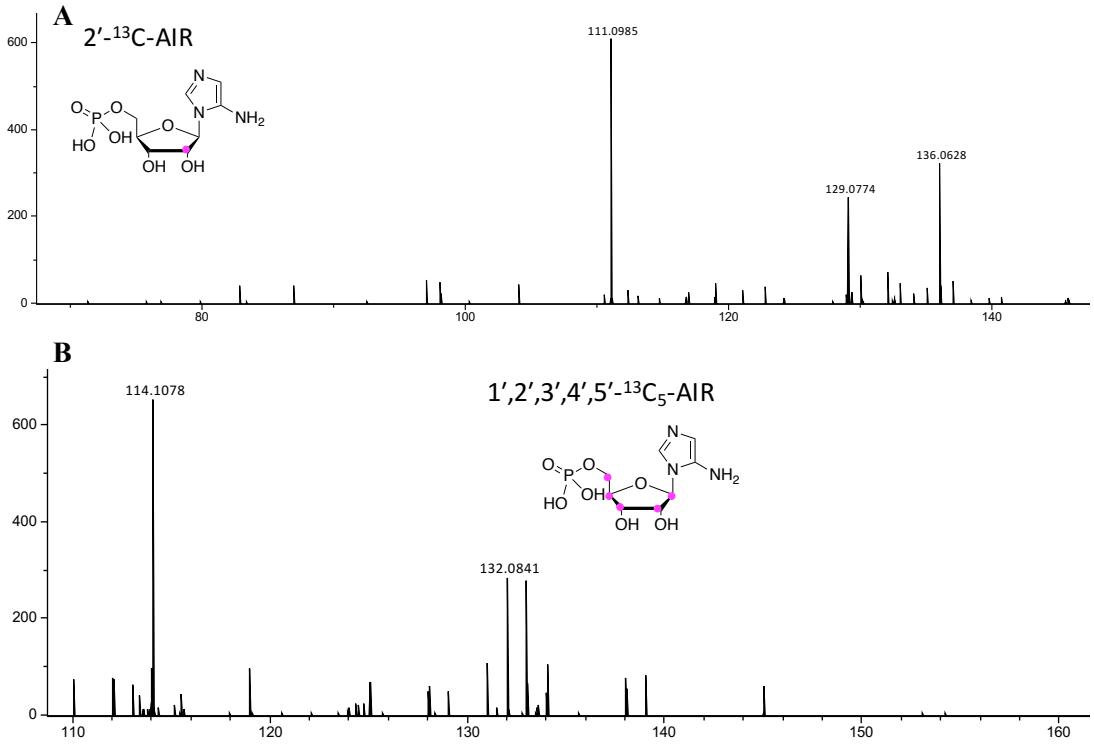


Figure 5.15 Mass fragmentation of the shunt product (204.1053) in ThiC catalyzed reaction when substrate is (A) 2′-^{13}C-AIR; (B) 1′,2′,3′,4′,5′-$^{13}C_5$-AIR; (C) 4′-^{2}H-AIR.

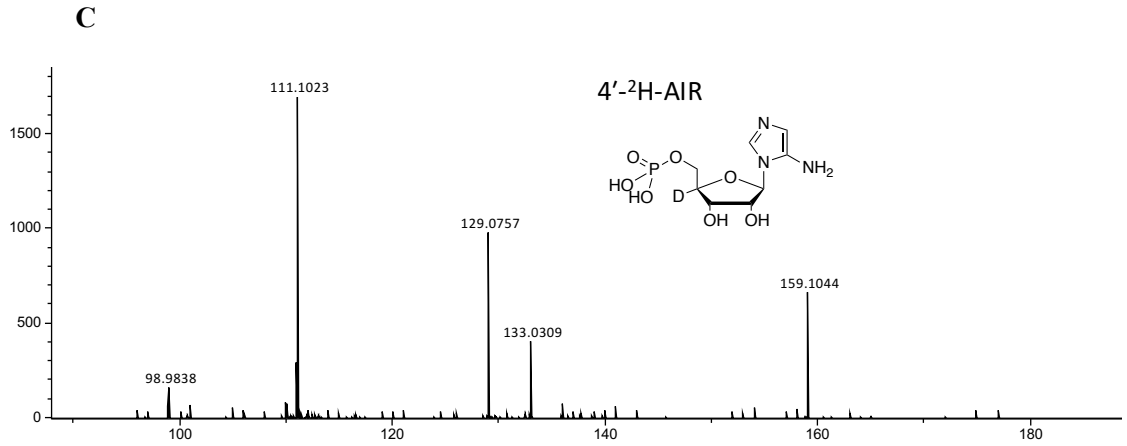


Figure 5.15 Mass fragmentation of the shunt product (204.1053) in ThiC catalyzed reaction when substrate is (A) 2′-^{13}C-AIR; (B) 1′,2′,3′,4′,5′-$^{13}C_5$-AIR; (C) 4′-^{2}H-AIR.

5.3. Conclusion

By using high resolution LC-MS, isotope labeling AIR, SAM and methionine, a four-carbon fragment and methionine adduct was detected. Since the adduct does not have a structure like compounds **5-22**. It's possible that the shunt product has structures like compounds **5-21**. From the isotope labeling studies, we learn that this four-carbons fragment forms after 5′-dA abstracts a hydrogen atom from the C5′ of the AIR but before the second hydrogen abstraction from C4′ of the AIR. Also, it forms after the C1′ release as formate and before imidazole recombines with the sugar backbone or the imidazole gets hydrolyzed after recombines with the sugar backbone. If it's not the latter case, it implies that C1′ leaves before the imidazole recombines during the ThiC catalyzed AIR rearrangement.

5.4. Experimental methods

5.4.1. Overexpression and purification of AtThiC

ΔN71-AtThiC in pET-28 plasmid was transformed into *E. coli* BL21(DE3) cells. ThiC was over expressed in *E. coli* BL21 DE3 and was co-expressed with a plasmid encoding the *suf* operon. A 15 mL overnight culture was grown in LB medium in the presence of 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. This was then added to 1.5 L of LB medium containing 40 μg/mL kanamycin and 30 μg/mL chloramphenicol. The cultures were incubated at 37°C with shaking (180 rpm) until the OD600 reached 0.45. The culture was then incubated at 4°C without shaking for 1 hour. Expression was induced by the addition of 100 μM IPTG followed by ferrous ammonium sulfate (120 mg) and cysteine (120 mg) and the culture was grown at 15°C with shaking (110 rpm) overnight. The cells were harvested by centrifugation and stored in liquid nitrogen until further use. All steps for protein purification were carried out in an anaerobic chamber (COY laboratories). Cell pellets (~12 g) were thawed and resuspended in 30 mL lysis buffer (100 mM Tris-HCl, pH 7.5) in the presence of lysozyme (0.2 mg/mL) and benzonase (100 units). This mixture was cooled in an ice-bath with continuous stirring for 1-2 hrs. The suspension was then sonicated to complete cell lysis for 5 × 120 s using a Misonix Sonicator XL-2000 (setting = 9). Cell debris was removed by centrifugation, and the lysate was loaded onto a Histrap column pre-equilibrated with lysis buffer. The column was washed with 5-10 column volumes of wash buffer (100 mM Tris-HCl, 30 mM Imidazole, pH 7.5). The CysS protein was then eluted using elution buffer (100 mM Tris-HCl, 300 mM NaCl, 250 mM imidazole, pH

7.5). The brown-colored fractions, containing the desired protein, were pooled and buffer exchanged to final buffer (100 mM Tris-HCl, pH 7.5 at 25 °C, 30% glycerol) using Econo-Pac 10DG size exclusion chromatography. The purified ThiC was divided into aliquots in eppendorf tubes, frozen in liquid nitrogen, and stored at −80 °C, submerged in liquid nitrogen, until future use. Protein concentration was measured by the absorbance at 280 nm (A280) with an extinction coefficient calculated using the ProtParam tool of the ExPASy proteomics server.

5.4.2. Enzymatic synthesis of AIR from AIRs:

A typical reaction mixture consisted of 25 µL 100 mM AIRs, 10 µL 400 mM ATP, 5 µL 1 M $MgCl_2$, 900 µL 25 mM KPi (pH 7.5) and 100 µL of the AIRs kinase stock.

The reaction mixture was allowed to incubate at room temperature for 1 hr and was lyophilized till dry. The residue was transferred into the glove box and resuspended in degassed water to a final volume of 100 µL (final AIR concentration of ~25 mM). This AIR stock was directly used for ThiC reactions.

5.4.3. Enzymatic assay of ThiC reaction

A typical enzymatic reaction was performed in 100 mM phosphate buffer, pH 7.5 containing 250 µM ThiC, 1 Mm AIR, 1.2 mM SAM and 5 mM dithionite. The reaction was incubated at room temperature anaerobically for 2 hours. EcFld (5 µM) /EcFpr (2 µM), NADPH (200 µM) reducing system anaerobically in dark. For HPLC, LC-MS analysis, the enzymatic reaction was heat quenched at 90 °C for two minutes and then centrifuged and filtered through a 10kDa cut-off filter.

5.4.4. Alkaline phosphate treatment of ThiC reaction

Alkaline phosphatase 1µL (1000 units) was added to 100 µl to 200 µl of ThiC enzymatic reaction. The mixture was incubated at 37°C for 30 minutes and then quenched through 10kDa cut-off filter before it was analyzed by HPLC or LC-MS.

5.4.5. HPLC parameters for the analysis of the ThiC reaction mixture

HPLC analysis was performed on a ZORBAX Eclipse XDB-C18 column (15cm 4.6mm, 5µm particles, Agilent Technologies, flow rate1 mL/min). Typical injection volumes were in the range of 10–80µL. Data were processed using ChemStation ver. B.04.01 SP1 (Agilent Technologies).

HPLC conditions: (Flow rate 1mL/min)

A. Water.

B. 100 mM potassium

acetate, pH 6.6. C. Methanol.

phosphate

buffer, pH

6.6/10 mM ammonium

HPLC method:

0min—100% B, 5min—100% B, 14min—7% A 70% B 23% C, 25min—25% A 0% B 75% C, 28min—25% A 0% B 75% C, 32min—100% B, 36min—100% B.

5.4.6. LC-MS parameters for the analysis of the ThiC reaction mixture

LC-ESI-TOF-MS on small molecules was performed using an Agilent 1260 HPLC system equipped with a binary pump and a 1200 series diode array detector

followed by a MicroToF-Q II mass spectrometer (Bruker Daltonics) using an ESI source in positive mode. The analysis was performed on a ZORBAX Eclipse XDB-C18 column (15 cm x 4.6 mm, 5 μm particles, Agilent).

LC conditions: A- 5 mM Ammonium acetate buffer, pH 6.6, B- 75 % Methanol and 25 % Water.

LC method: 0 min – 100% A, 2 min – 100% A, 4 min – 80% A 20% B, 27 min – 100% B, 29 min – 100% B, 30 min – 100% A, 40 min – 100% A.

6. SUMMARY AND OUTLOOK

6.1. Summary

Here, we have explored unusual reactions catalyzed by radical SAM enzymes. In cystobactamid biosynthesis, CysS utilizes B_{12}, SAM and [4Fe-4S] to perform sequential methylations on an sp^3-hybridized carbon. In anaerobic heme degradation pathway, ChuW not only methylates the porphyrin ring but also cleaves the ring to release the iron. In thiamin biosynthesis pathway, ThiC catalyzes the most complicated rearrangement from AIR to HMP-P. In the first two cases, classic probes (radical trap, isotope labeling compounds) were designed to trap the unstable intermediates to study the detailed mechanism or conformations in the enzyme active site. For the last case, because of the narrow substrates tolerance of the ThiC and the lack of chromophore in the substrate, it's harder to use substrate analogs to study the mechanism. Here, we are able to trap a four-carbon fragmentation from AIR by methionine using unbiased LC-MS/MS analysis and isotope labeling approaches. These give us a better insight into how nature uses complex protein architecture and inherent chemistry property of substrates to perform novel reactions.

6.2. Outlook

6.2.1. B_{12} dependent radical SAM methyltransferase-CysS

6.2.1.1. Crystallography

Currently, there is no crystal structure that is published for the B_{12} dependent radical SAM methyltransferase. However, there is a crystal structure for a B_{12}-dependent

radical SAM enzyme, OxsB, which performs a ring contraction reaction to generate oxetanocin A. Although no substrate or product is bound to the crystal, cobalamin, SAM, [4Fe-4S] cluster are found in the crystal, which gives us insight into how a B_{12}-dependent radical SAM transferase could position the SAM and Cbl to perform both racial reaction and methyl transfer. Bridwell-Rabb was able to obtain two conformations of SAM binding to the unique Fe of the [4Fe-4S] cluster.[22] In one of the conformation, the S atom of SAM is 3.0 Å to the unique Fe, which is in the range (3.4 Å) that is permitted for reductive cleavage of SAM while the methyl group of the SAM is 8.8 Å from Co of Cbl. In another conformation, the S atom of SAM is 6.5 Å to the unique Fe while the methyl group of the SAM is 5.8 Å from Co of Cbl, which favors the methyl transfer reaction to the Cbl. In this case, it seems SAM can both initiate the radical chemistry and perform the methyl transfer in the same active site. Nevertheless, OxsB is not a methyltransferase. It's not clear why SAM adopt two conformations to perform the ring contraction and why Cbl is needed. CysS has a sequence similarity around 16% with OxsB and mostly align in the Cbl-binding domain and SAM radical domain. Compared to the OxsB, CysS lacks the N-terminal domain of OxsB and has great variation in the C-terminal helix bundle of OxsB, which is proposed for the conformational change upon the binding of the substrate. We have made several constructs and substrate analogs for the crystallography of CysS and the study is still in progress. It will be interesting to see whether CysS adopts similar conformations for SAM to achieve the unusual methylation with Cbl.

6.2.1.2. The evolution of radical SAM enzymes

Currently, the radical SAM superfamily contains 113,000 unique sequences that represent more than 65 distinct reactions across all domains of life. The catalytic activities of radical SAM enzymes are usually low compared to other highly efficient enzymes. SAM is not always used as a cofactor instead is consumed as a substrate for many reactions catalyzed by radical SAM enzymes. Both [4Fe-4S] cluster and SAM are not complicated cofactors in the living organism. Therefore, the radical SAM enzymes are considered as ancient enzymes instead of well evolved modern enzymes, especially for their capability to mediate radical reactions under anaerobic condition.

Adenosylcobalamin (AdoCbl)-dependent enzymes also catalyze unusual rearrangement or elimination reactions through radical chemistry. The catalytic cycle starts with the homolytic cleavage of the Co–C bond to generated the 5′-dA radical, which abstracts a hydrogen atom from a substrate. The substrate radical can undergo rearrangement and subsequent hydrogen exchange to yield the final product.[90] Most of the radical SAM enzymes also use 5′-dA radicals to initiate the reaction. However, The C–Co bond of AdoCbl is around 30 kcal/mol while the 5′-C-S bond of SAM is around 60 kcal/mol. Also, in order to cleave SAM, there must be a one-electron reduction step. Considered SAM is not as a chemically elegant molecule as adenosylcobalamin, SAM is regarded as "a poor man's adenosylcobalamin.[91] However, in terms of the number of the enzymes and reaction diversity, the radical SAM enzymes far exceed of the AdoCbl-dependent enzymes. There are only 12 groups of AdoCbl-dependent enzymes have so far been identified, which include skeleton mutase, eliminase and aminomutase.

Furthermore, it takes nearly 30 enzymatic steps to biosynthesize cobalamin while SAM only requires one step by methionine adenosyltransferase using methionine and ATP. From these aspects, it's possible that AdoCbl-dependent enzymes evolve from radical SAM enzymes.

Cobalamin-dependent radical SAM enzymes are a large class of enzymes that use both radical SAM machinery and cobalamin to catalyze unusual reactions, such as inactive carbon methylation, phosphorous atoms methylation, ring contraction. From the crystal structure of OxsB, we know that SAM bound [4Fe-4S] cluster and Cbl are just 4 Å apart, which indicates that it's quite conceivable that same substrate can bind to a single position and react with two cofactors in a single active site. This active site combines both SAM radical core unit and Cbl-binding Rossmann domain. It's proposed that OxsB can be considered as a radical SAM enzyme that has been adapted to bind Cbl.[92] It's likely that cobalamin-dependent radical SAM enzyme is an ancient ancestry of the AdoCbl-dependent enzyme.

From the mechanistic studies of CysS, we learn that the enzyme prefers Cbl(II) to start with and 5′-dA radical can be generated by reductive cleavage of radical SAM enzyme. With a close vicinity of Cbl(II) and a 5′-dA radical, it's interesting to speculate the possibility of generating the AdoCbl, which could be a random event for the first appearance of the sophisticated cofactor AdoCbl.

6.2.2. ChuW in anaerobic heme degradation pathway

ChuW has high sequence similarity with HemN. HemN is a radical SAM enzyme that catalyzes the oxidative decarboxylation of the propionate side chains of rings A and B of coproporphyrinogen III to form protoporphyrinogen IX. It is shown that EPR can be used to detect a comparatively stable delocalized substrate radical intermediate.[93] Here for ChuW, we are able to use a cyclopropyl heme analog to trap a radical intermediate on the porphyrin ring by forming stable porphyrin compound and SAM adducts. It's conceivable that we can directly use EPR to detect the possible radical intermediate in the ChuW catalyzed cyclopropyl heme analog reaction.

ChuW is assigned as the class C radical SAM enzyme. NosN, another enzyme in this family, catalyzes both carbon transfer from SAM and formation of the ester linkage between Glu6 and the appended C4 carbon unit of 3,4-dimethylindolic acid (DMIA) motif in nosiheptide biosynthesis.[38] It is shown that NosN can cleave an allyl analog of SAM (allyl-SAM) efficiently and the resulting 5′-dA radical can be captured by the olefin moieties of allyl-SAM or 5′-allylthioadenosine (ATA).[94] It's possible that ChuW can also take allyl-SAM as a substrate and the resulting product may provide insight of whether 5′-dA radical abstracts a hydrogen atom from SAM or methylthioadenosine, and the orientation of two SAM molecules in the active site. It will also be useful to get a crystal structure because currently there is no crystal structure available for this kind of enzymes.

6.2.3. HMP-P synthase ThiC in thiamin pyrimidine biosynthesis

6.2.3.1. Formation of the four-carbon fragment with methionine

From the structure characterization of the four-carbon fragment, we know that all the oxygen is lost in this shunt product. It will be useful to know whether the reducing agent- dithionite used for the ThiC reaction is strong enough to reduce the leaking intermediate to form the four-carbon fragment or the two oxygen indeed is lost during the enzymatic reaction. If it's the latter case, the oxygen of the by-product carbon monoxide in the ThiC catalyzed reaction should come from water instead of oxygen from the AIR. We can try to obtain the isotope pattern of oxygen of carbon monoxide in ThiC reaction under D_2O condition by GC-MS.

6.2.3.2. The function of the mononuclear metal center in ThiC

Compare to other classic radical SAM enzymes, ThiC not only has a different binding mode for SAM but also have an extra Fe/Zn binding site. It will be interesting to know whether Fe is the actual metal that function in ThiC and whether it has a redox role in catalyzing the reaction. Also, the ThiC crystal structure has shown the amino and carboxylate group of methionine are coordinated with the metal. The metal may also play a role in trapping the four-carbon fragment with methionine. Mutants study of the ligands (H417A, H481A) on the effect of the formation of the four-carbon fragment with methionine will give some hints about how the four-carbon fragment forms.

6.2.3.3. ThiC mutants' studies

Besides H417A and H481A in *Caulobacter crescentus* ThiC (CcThiC), other mutants (C474S, Y277F, M248L in CcThiC, D383N in AtThiC) show the 5′-dA radical

abstracts a hydrogen atom from AIR, which suggests the radical chemistry is initiated. Potential intermediate can be trapped by radical probes for these mutants.

6.2.3.4. Substrate analogs studies

In the proposed mechanism for ThiC catalyzed rearrangement, there is a dehydration reaction induced by the radical, which is similar to the case in ribonucleotide reductase (RNR) for the step **5-14** to **5-16** in Figure 5.2. There have been extensive substrate analogs studies for RNR. Similar strategy can also be applied to study the mechanism of ThiC.

REFERENCES

(1) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E., *Nucleic acids research* **2001,** *29*, 1097-1106.

(2) Atkinson, J. K.; Ingold, K. U., *Biochemistry* **1993,** *32*, 9209-9214.

(3) Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M., *Chem Rev* **2014,** *114*, 4229-4317.

(4) Chatterjee, A.; Li, Y.; Zhang, Y.; Grove, T. L.; Lee, M.; Krebs, C.; Booker, S. J.; Begley, T. P.; Ealick, S. E., *Nat Chem Biol* **2008,** *4*, 758-765.

(5) Mehta, A. P.; Abdelwahed, S. H.; Fenwick, M. K.; Hazra, A. B.; Taga, M. E.; Zhang, Y.; Ealick, S. E.; Begley, T. P., *Journal of the American Chemical Society* **2015,** *137*, 10444-10447.

(6) Mahanta, N.; Fedoseyenko, D.; Dairi, T.; Begley, T. P., *Journal of the American Chemical Society* **2013,** *135*, 15318-15321.

(7) Joshi, S.; Mahanta, N.; Fedoseyenko, D.; Williams, H.; Begley, T. P., *Journal of the American Chemical Society* **2017,** *139*, 10952-10955.

(8) Philmus, B.; Decamps, L.; Berteau, O.; Begley, T. P., *Journal of the American Chemical Society* **2015,** *137*, 5406-5413.

(9) Mehta, A. P.; Hanes, J. W.; Abdelwahed, S. H.; Hilmey, D. G.; Hanzelmann, P.; Begley, T. P., *Biochemistry* **2013,** *52*, 1134-1136.

(10) Hover, B. M.; Tonthat, N. K.; Schumacher, M. A.; Yokoyama, K., *P Natl Acad Sci USA* **2015,** *112*, 6347-6352.

(11) Akiva, E.; Brown, S.; Almonacid, D. E.; Barber, A. E., 2nd; Custer, A. F.; Hicks, M. A.; Huang, C. C.; Lauck, F.; Mashiyama, S. T.; Meng, E. C.; Mischel, D.; Morris, J. H.; Ojha, S.; Schnoes, A. M.; Stryke, D.; Yunes, J. M.; Ferrin, T. E.; Holliday, G. L.; Babbitt, P. C., *Nucleic acids research* **2014,** *42*, D521-30.

(12) Kamat, S. S.; Williams, H. J.; Dangott, L. J.; Chakrabarti, M.; Raushel, F. M., *Nature* **2013,** *497*, 132-136.

(13) Dowling, D. P.; McCarty, R. M.; Young, A. P.; Bruender, N. A.; Bandarian, V.; Drennan, C. L., *Abstr Pap Am Chem S* **2013,** *246*.

(14) Dowling, D. P.; Vey, J. L.; Croft, A. K.; Drennan, C. L., *Bba-Proteins Proteom* **2012,** *1824*, 1178-1195.

(15) Fenwick, M. K.; Mehta, A. P.; Zhang, Y.; Abdelwahed, S. H.; Begley, T. P.; Ealick, S. E., *Nature communications* **2015,** *6*.

(16) Dong, M.; Kathiresan, V.; Fenwick, M. K.; Torelli, A. T.; Zhang, Y.; Caranto, J. D.; Dzikovski, B.; Sharma, A.; Lancaster, K. M.; Freed, J. H.; Ealick, S. E.; Hoffman, B. M.; Lin, H. N., *Science* **2018,** *359*, 1247-1250.

(17) Grillo, M. A.; Colombatto, S., *Amino acids* **2008,** *34*, 187-93.

(18) Woodard, R. W.; Tsai, M. D.; Floss, H. G.; Crooks, P. A.; Coward, J. K., *J Biol Chem* **1980,** *255*, 9124-9127.

(19) Zhang, Q.; van der Donk, W. A.; Liu, W., *Accounts Chem Res* **2012,** *45*, 555-564.

(20) Bauerle, M. R.; Schwalm, E. L.; Booker, S. J., *J Biol Chem* **2015,** *290*, 3995-4002.

(21) Wang, S. C., *Nat Prod Rep* **2018,** *35*, 707-720.

(22) Bridwell-Rabb, J.; Zhong, A. S.; Sun, H. G.; Drennan, C. L.; Liu, H. W., *Nature* **2017,** *544*, 322-U321.

(23) van der Donk, W. A., *J Org Chem* **2006,** *71*, 9561-9571.

(24) Lanz, N. D.; Blaszczyk, A. J.; McCarthy, E. L.; Wang, B.; Wang, R. X.; Jones, B. S.; Booker, S. J., *Biochemistry* **2018,** *57*, 1475-1490.

(25) Werner, W. J.; Allen, K. D.; Hu, K. F.; Helms, G. L.; Chen, B. S.; Wang, S. C., *Biochemistry* **2011,** *50*, 8986-8988.

(26) Kim, H. J.; Liu, Y. N.; McCarty, R. M.; Liu, H. W., *Journal of the American Chemical Society* **2017,** *139*, 16084-16087.

(27) McLaughlin, M. I.; van der Donk, W. A., *Biochemistry* **2018**.

(28) Sato, S.; Kudo, F.; Kuzuyama, T.; Hammerschmidt, F.; Eguchi, T., *Biochemistry* **2018**.

(29) Wang, B.; Blaszczyk, A. J.; Knox, H. L.; Zhou, S.; Blaesi, E. J.; Krebs, C.; Wang, R. X.; Booker, S. J., *Biochemistry* **2018,** *57*, 4972-4984.

(30) Marous, D. R.; Lloyd, E. P.; Buller, A. R.; Moshos, K. A.; Grove, T. L.; Blaszczyk, A. J.; Booker, S. J.; Townsend, C. A., *P Natl Acad Sci USA* **2015,** *112*, 10354-10358.

(31) Benjdia, A.; Pierre, S.; Gherasim, C.; Guillot, A.; Carmona, M.; Amara, P.; Banerjee, R.; Berteau, O., *Nature communications* **2015,** *6*, 8377.

(32) Blaszczyk, A. J.; Wang, B.; Silakov, A.; Ho, J. V.; Booker, S. J., *J Biol Chem* **2017,** *292*, 15456-15467.

(33) Parent, A.; Guillot, A.; Benjdia, A.; Chartier, G.; Leprince, J.; Berteau, O., *Journal of the American Chemical Society* **2016,** *138*, 15515-15518.

(34) Freeman, M. F.; Helf, M. J.; Bhushan, A.; Morinaka, B. I.; Piel, J., *Nat Chem* **2017,** *9*, 387-395.

(35) Wang, Y.; Schnell, B.; Baumann, S.; Muller, R.; Begley, T. P., *Journal of the American Chemical Society* **2017,** *139*, 1742-1745.

(36) Layer, G.; Moser, J.; Heinz, D. W.; Jahn, D.; Schubert, W. D., *The EMBO journal* **2003,** *22*, 6214-24.

(37) Ding, W.; Wu, Y. J.; Ji, X. J.; Qianzhu, H. C.; Chen, F.; Deng, Z. X.; Yu, Y.; Zhang, Q., *Chemical communications* **2017,** *53*, 5235-5238.

(38) LaMattina, J. W.; Wang, B.; Badding, E. D.; Gadsby, L. K.; Grove, T. L.; Booker, S. J., *Journal of the American Chemical Society* **2017,** *139*, 17438-17445.

(39) Hiratsuka, T.; Suzuki, H.; Kariya, R.; Seo, T.; Minami, A.; Oikawa, H., *Angew Chem Int Edit* **2014,** *53*, 5423-5426.

(40) Huang, W.; Xu, H.; Li, Y.; Zhang, F.; Chen, X. Y.; He, Q. L.; Igarashi, Y.; Tang, G. L., *Journal of the American Chemical Society* **2012,** *134*, 8831-8840.

(41) Jin, W. B.; Wu, S.; Jian, X. H.; Yuan, H.; Tang, G. L., *Nature communications* **2018,** *9*.

(42) Mahanta, N.; Zhang, Z. A.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A., *Journal of the American Chemical Society* **2017,** *139*, 4310-4313.

(43) Hein, S.; von Irmer, J.; Gallei, M.; Meusinger, R.; Simon, J., *Bba-Bioenergetics* **2018,** *1859*, 300-308.

(44) Maier, G.; Pfriem, S.; Schafer, U.; Matusch, R., *Angewandte Chemie-International Edition in English* **1978,** *17*, 520-521.

(45) Bisel, P.; Al-Momani, L.; Muller, M., *Org Biomol Chem* **2008,** *6*, 2655-2665.

(46) Nakanishi, K.; Habaguchi, K., *Journal of the American Chemical Society* **1971,** *93*, 3546-3547.

(47) Mujumdar, R. B.; Rao, A. V. R.; Rathi, S. S.; Venkataraman, K., *Tetrahedron Lett* **1975**, 867-868.

(48) Rezanka, T.; Kolouchova, I.; Cejkova, A.; Sigler, K., *Appl Microbiol Biot* **2012,** *95*, 1371-1376.

(49) Kitanishi, K.; Cracan, V.; Banerjee, R., *J Biol Chem* **2015,** *290*, 20466-76.

(50) Skiba, M. A.; Sikkema, A. P.; Moss, N. A.; Lowell, A. N.; Su, M.; Sturgis, R. M.; Gerwick, L.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L., *ACS chemical biology* **2018,** *13*, 1640-1650.

(51) Baumann, S.; Herrmann, J.; Raju, R.; Steinmetz, H.; Mohr, K. I.; Huttel, S.; Harmrolfs, K.; Stadler, M.; Muller, R., *Angewandte Chemie* **2014,** *53*, 14605-9.

(52) Hanzelmann, P.; Hernandez, H. L.; Menzel, C.; Garcia-Serres, R.; Huynh, B. H.; Johnson, M. K.; Mendel, R. R.; Schindelin, H., *J Biol Chem* **2004,** *279*, 34721-32.

(53) Salnikov, D. S.; Silaghi-Dumitrescu, R.; Makarov, S. V.; van Eldik, R.; Boss, G. R., *Dalton T* **2011,** *40*, 9831-9834.

(54) Mosimann, H.; Kräutler, B., *Angewandte Chemie International Edition* **2000,** *39*, 393-395.

(55) Wang, Y.; Schnell, B.; Muller, R.; Begley, T. P., *Methods Enzymol* **2018,** *606*, 199-216.

(56) Lohman, J. R.; Huang, S. X.; Horsman, G. P.; Dilfer, P. E.; Huang, T. T.; Chen, Y. H.; Wendt-Pienkowski, E.; Shen, B., *Mol Biosyst* **2013,** *9*, 478-491.

(57) Watanabe, K.; Hotta, K.; Nakaya, M.; Praseuth, A. P.; Wang, C. C.; Inada, D.; Takahashi, K.; Fukushi, E.; Oguri, H.; Oikawa, H., *Journal of the American Chemical Society* **2009,** *131*, 9347-53.

(58) Kuzuyama, T.; Seki, T.; Dairi, T.; Hidaka, T.; Seto, H., *J Antibiot* **1995,** *48*, 1191-1193.

(59) Hu, K. F.; Werner, W. J.; Allen, K. D.; Wang, S. C., *Magn Reson Chem* **2015,** *53*, 267-272.

(60) Huttel, S.; Testolin, G.; Herrmann, J.; Planke, T.; Gille, F.; Moreno, M.; Stadler, M.; Bronstrup, M.; Kirschning, A.; Muller, R., *Angewandte Chemie* **2017,** *56*, 12760-12764.

(61) Studer, A.; Steen, H., *Chem-Eur J* **1999,** *5*, 759-773.

(62) Parry, R. J., *Journal of the American Chemical Society* **1977,** *99*, 6464-6466.

(63) Fugate, C. J.; Jarrett, J. T., *Bba-Proteins Proteom* **2012,** *1824*, 1213-1222.

(64) Douglas, P.; Kriek, M.; Bryant, P.; Roach, P. L., *Angew Chem Int Edit* **2006,** *45*, 5197-5199.

(65) Lanz, N. D.; Pandelia, M. E.; Kakar, E. S.; Lee, K. H.; Krebs, C.; Booker, S. J., *Biochemistry* **2014,** *53*, 4557-4572.

(66) Huang, W. L.; Wilks, A., *Annual Review of Biochemistry, Vol 86* **2017,** *86*, 799-823.

(67) Fischbach, M. A.; Lin, H. N.; Liu, D. R.; Walsh, C. T., *Nat Chem Biol* **2006,** *2*, 132-138.

(68) Andrews, S. C.; Robinson, A. K.; Rodriguez-Quinones, F., *Fems Microbiol Rev* **2003,** *27*, 215-237.

(69) Kumar, S.; Bandyopadhyay, U., *Toxicol Lett* **2005,** *157*, 175-188.

(70) Lyles, K. V.; Eichenbaum, Z., *Front Cell Infect Mi* **2018,** *8*.

(71) Sheldon, J. R.; Heinrichs, D. E., *Fems Microbiol Rev* **2015,** *39*, 592-630.

(72) de Montellano, P. R. O., *Curr Opin Chem Biol* **2000,** *4*, 221-227.

(73) Docherty, J. C.; Schacter, B. A.; Firneisz, G. D.; Brown, S. B., *J Biol Chem* **1984,** *259*, 3066-3069.

(74) Docherty, J. C.; Firneisz, G. D.; Schacter, B. A., *Arch Biochem Biophys* **1984,** *235*, 657-664.

(75) Tenhunen, R.; Marver, H. S.; Schmid, R., *J Biol Chem* **1969,** *244*, 6388-&.

(76) Wyckoff, E. E.; Schmitt, M.; Wilks, A.; Payne, S. M., *J Bacteriol* **2004,** *186*, 4142-4151.

(77) Septer, A. N.; Wang, Y. L.; Ruby, E. G.; Stabb, E. V.; Dunn, A. K., *Environmental microbiology* **2011,** *13*, 2855-2864.

(78) LaMattina, J. W.; Nix, D. B.; Lanzilotta, W. N., *P Natl Acad Sci USA* **2016,** *113*, 12138-12143.

(79) Beckwith, A. L. J.; Bowry, V. W., *Journal of the American Chemical Society* **1994,** *116*, 2710-2716.

(80) Vicente, M. G., *Curr Med Chem Anticancer Agents* **2001,** *1*, 175-94.

(81) Kozyrev, A. N.; Alderfer, J. L.; Robinson, B. C., *Tetrahedron* **2003,** *59*, 499-504.

(82) Gouterman, M., *J Mol Spectrosc* **1961,** *6*, 138-&.

(83) Iwig, D. F.; Booker, S. J., *Biochemistry* **2004,** *43*, 13496-13509.

(84) Torpey, J. W.; Demontellano, P. R. O., *J Org Chem* **1995,** *60*, 2195-2199.

(85) Jurgenson, C. T.; Begley, T. P.; Ealick, S. E., *Annu Rev Biochem* **2009,** *78*, 569-603.

(86) Begley, T. P.; Ealick, S. E.; McLafferty, F. W., *Biochem Soc T* **2012,** *40*, 555-560.

(87) Lawhorn, B. G.; Mehl, R. A.; Begley, T. P., *Org Biomol Chem* **2004,** *2*, 2538-2546.

(88) Chatterjee, A.; Hazra, A. B.; Abdelwahed, S.; Hilmey, D. G.; Begley, T. P., *Angew Chem Int Edit* **2010,** *49*, 8653-8656.

(89) Gagnon, D. M.; Stich, T. A.; Mehta, A. P.; Abdelwahed, S. H.; Begley, T. P.; Britt, R. D., *Journal of the American Chemical Society* **2018,** *140*, 12798-12807.

(90) Marsh, E. N. G.; Melendez, G. D. R., *Bba-Proteins Proteom* **2012,** *1824*, 1154-1164.

(91) Frey, P. A., *Faseb J* **1993,** *7*, 662-670.

(92) Bridwell-Rabb, J.; Grell, T. A. J.; Drennan, C. L., *Annual Review of Biochemistry, Vol 87* **2018,** *87*, 555-584.

(93) Layer, G.; Pierik, A. J.; Trost, M.; Rigby, S. E.; Leech, H. K.; Grage, K.; Breckau, D.; Astner, I.; Jansch, L.; Heathcote, P.; Warren, M. J.; Heinz, D. W.; Jahn, D., *J Biol Chem* **2006,** *281*, 15727-34.

(94) Ji, X. J.; Mandalapu, D.; Cheng, J. D.; Ding, W.; Zhang, Q., *Angew Chem Int Edit* **2018,** *57*, 6601-6604.